



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 121831

TO: Jon Ashen
Location: REM-2B02/2C18
Art Unit: 1635
Friday, May 14, 2004

Case Serial Number: 09/673994

From: Toby Port
Location: Biotech-Chem Library
Remsen 1A59
Phone: 571-272-2523

toby.port@uspto.gov

Search Notes

Dear Examiner Ashen,

Here are the results of your search.
Please feel free to contact me if you have any questions.

Toby Port



modulates expression of the gene. 2 Other isoforms (see AAY32376-77) of CNREB-2 have been identified, the shortest isoform showing strongest binding activity to the CNRE of the renin promoter. CNREB-2 shares no significant homology to any other known gene or protein. It is a novel zinc finger transcription factor. Expression vectors, isolated nucleic acid molecules and CNREB-2 polypeptides are claimed. The invention also relates to CNREB-1 polynucleotides (see AAZ5040-43), polypeptides (see AAY32374) and inhibitors, and their use in claimed methods for decreasing or increasing renin expression, and for determining the level of CNREB-1 expression, an increase in expression being indicative of an increased susceptibility to development of a renin-angiotensin system mediated disorder

XX Sequence 3847 BP; 805 A; 1008 C; 1171 G; 852 T; 0 U; 11 Other;

Query Match 99.7%; Score 3836; DB 3; Length 3847;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 3847; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	GGAGGTGTAGCTTTATTCATCAAGATCAGTCAACGTGATGAAGAAACGGGAGCTTCG	60
DB	1	GGAGGTGTAGCTTTATTCATCAAGATCAGTCAACGTGATGAAGAAACGGGAGCTTCG	60
QY	61	GTATTGTGTCATCCATTGTGGAAGAGGTGAGTCTAAAGTTGGGAGGAGNGNTGGTNTA	120
DB	61	GTATTGTGTCATCCATTGTGGAAGAGGTGAGTCTAAAGTTGGGAGGAGNGNTGGTNTA	120
QY	121	AAACAGANAAATTTTTTGGAGTGTGAGCAGAGCTTGGGAGTTAGATTCACTGAANACAA	180
DB	121	AAACAGANAAATTTTTTGGAGTGTGAGCAGAGCTTGGGAGTTAGATTCACTGAANACAA	180
QY	181	TTAGTGACTTTGATTAAGANGNGGTTNTTCTTCAATTTCCAGATGTAGCTGGGCTTC	240
DB	181	TTAGTGACTTTGATTAAGANGNGGTTNTTCTTCAATTTCCAGATGTAGCTGGGCTTC	240
QY	241	AGAAATNACNGTTTTTTTTTGTGATTAACCTTGCCCATCTTTGTCTGCTTCTTCTTG	300
DB	241	AGAAATNACNGTTTTTTTTTGTGATTAACCTTGCCCATCTTTGTCTGCTTCTTCTTG	300
QY	301	TATTAGCAAGCAAGCTTTCTTGTGTACTTTTGGTTTCAAAACCCATGGCTAACTGCCT	360
DB	301	TATTAGCAAGCAAGCTTTCTTGTGTACTTTTGGTTTCAAAACCCATGGCTAACTGCCT	360
QY	361	TAACTTTTGTAGCTGCTGGATCAGTGCCTGGCGGACACATCTCTGGGATGGTACAG	420
DB	361	TAACTTTTGTAGCTGCTGGATCAGTGCCTGGCGGACACATCTCTGGGATGGTACAG	420
QY	421	GGTGTGACATCATGCTGATCATCATTTTGAAGTTCAGTCCGCAACACAGTAGAGCTTC	480
DB	421	GGTGTGACATCATGCTGATCATCATTTTGAAGTTCAGTCCGCAACACAGTAGAGCTTC	480
QY	481	AGGAGAGAGTGGCAAGAGAACTGGAGAGGACCTGTCCAGGGGGGCTGAAGACA	540
DB	481	AGGAGAGAGTGGCAAGAGAACTGGAGAGGACCTGTCCAGGGGGGCTGAAGACA	540
QY	541	CCGGACACGTGACAAAGAGAGCCAGAGGATAGTGGTCAACCGGACCTGCCCAT	600
DB	541	CCGGACACGTGACAAAGAGAGCCAGAGGATAGTGGTCAACCGGACCTGCCCAT	600
QY	601	GTGACTGTTTCTGAGCCCTTGGACCCGAGACTGAGTTTGTCTTGTCTTTAGCCCTT	660
DB	601	GTGACTGTTTCTGAGCCCTTGGACCCGAGACTGAGTTTGTCTTGTCTTTAGCCCTT	660
QY	661	AGCAGTGGGTATGAGTGTGACAGGGGCTGGTGGCTTCTCAGCCATTACAAAGAGG	720
DB	661	AGCAGTGGGTATGAGTGTGACAGGGGCTGGTGGCTTCTCAGCCATTACAAAGAGG	720
QY	721	GCCCCCACCACCCCGGACGCTGGAGGCTCTGCTGCTCTTAAAGCTTCCTTA	780
DB	721	GCCCCCACCACCCCGGACGCTGGAGGCTCTGCTGCTCTTAAAGCTTCCTTA	780
QY	781	CTCTCTGGGCTCATGACTATCGTTCTGTGCTGTCTGTGTTGTGGAGGAG	840

DB	781	CTCTCTGGGCTCATGACTATCGGTTCTGTGCTGTGCTGTGTTGTGGAGGAG	840
QY	841	GACTGGTAGTCTGATTTTACTCTGTGAACACTTATTTAAGACATCTTTTATTG	900
DB	841	GACTGGTAGTCTGATTTTACTCTGTGAACACTTATTTAAGACATCTTTTATTG	900
QY	901	CGGCTCTGTGACCCCTAGCGGTTGCACCCGCTCTCTGTGTACACTTTCAAGACAC	960
DB	901	CGGCTCTGTGACCCCTAGCGGTTGCACCCGCTCTCTGTGTACACTTTCAAGACAC	960
QY	961	TTTTTTCAGACTAAAGGCCAAACAAAGCTAATCGTGTCTCATAGTGTATCTCTC	1020
DB	961	TTTTTTCAGACTAAAGGCCAAACAAAGCTAATCGTGTCTCATAGTGTATCTCTC	1020
QY	1021	CTACCTGGGCCCATATGTGGTGGCGGCTGTGTGCTCATGCTGTGTGTCGAGCC	1080
DB	1021	CTACCTGGGCCCATATGTGGTGGCGGCTGTGTGCTCATGCTGTGTGTCGAGCC	1080
QY	1081	TCTGGAGAGGGGGCAGTGAATGTGGAGCAGGAGCTGGGAGCTTTAACTCCTGAGTAA	1140
DB	1081	TCTGGAGAGGGGGCAGTGAATGTGGAGCAGGAGCTGGGAGCTTTAACTCCTGAGTAA	1140
QY	1141	GCCGCTGGAGGCCATCTGCGGCTGTTCGGCTTCAGGCACGACTTCTAGTTCGTATG	1200
DB	1141	GCCGCTGGAGGCCATCTGCGGCTGTTCGGCTTCAGGCACGACTTCTAGTTCGTATG	1200
QY	1201	GAGGGGGCAGGCTTACCACTTCTGCTGTGGTGGTGGTGGTGGTGGTGGTGGTGGT	1260
DB	1201	GAGGGGGCAGGCTTACCACTTCTGCTGTGGTGGTGGTGGTGGTGGTGGTGGTGGT	1260
QY	1261	GCTTGGCTAGCAGCGCATCAATCACTTCTGCTGTGGTGGTGGTGGTGGTGGTGGT	1320
DB	1261	GCTTGGCTAGCAGCGCATCAATCACTTCTGCTGTGGTGGTGGTGGTGGTGGTGGT	1320
QY	1321	ACAGCCCATGAATCAGCATCCCGTGTGCTGCTGGTGGCGGAGCGGAGCGGTAAG	1380
DB	1321	ACAGCCCATGAATCAGCATCCCGTGTGCTGCTGGTGGCGGAGCGGAGCGGTAAG	1380
QY	1381	GAGTACACAGCTGCTTCTGAGAGGGGGTAGGGGGCTCCATGAATGAAAGCGGGGG	1440
DB	1381	GAGTACACAGCTGCTTCTGAGAGGGGGTAGGGGGCTCCATGAATGAAAGCGGGGG	1440
QY	1441	GGCGGAGCAGCTGAGCTGATTCGCGGGCGGGGCGGGGCTGCCAGGGCCCGGACC	1500
DB	1441	GGCGGAGCAGCTGAGCTGATTCGCGGGCGGGGCGGGGCTGCCAGGGCCCGGACC	1500
QY	1501	GTGTATGGGGGGCGGTTCTGATGATCTTAAGAGCAAGAGCCGACCGGCGGCGGAA	1560
DB	1501	GTGTATGGGGGGCGGTTCTGATGATCTTAAGAGCAAGAGCCGACCGGCGGCGGAA	1560
QY	1561	GGGGCGGGCGGGGAGGCTCGGGCGGAGCGGCTGGGGCTGGAGCGGCTCAGATGCC	1620
DB	1561	GGGGCGGGCGGGGAGGCTCGGGCGGAGCGGCTGGGGCTGGAGCGGCTCAGATGCC	1620
QY	1621	CCGAAGGAAGCAAAAGCCATCCACAGCCCGTGAATGGAGGGGGTCAAAAGTCAAGG	1680
DB	1621	CCGAAGGAAGCAAAAGCCATCCACAGCCCGTGAATGGAGGGGGTCAAAAGTCAAGG	1680
QY	1681	AGGGGCTTGAAGCGGAGGAAACAGGGGTGGGTCAGTAGAGTGGGCTCAGGTACGGG	1740
DB	1681	AGGGGCTTGAAGCGGAGGAAACAGGGGTGGGTCAGTAGAGTGGGCTCAGGTACGGG	1740
QY	1741	GAGGGGACTCTCTCAGGGTTAGGGGGCGGATGATCTGGGATCTTCTGCTCTTACCA	1800
DB	1741	GAGGGGACTCTCTCAGGGTTAGGGGGCGGATGATCTGGGATCTTCTGCTCTTACCA	1800
QY	1801	TTAAAGGAACCTGAGGGTTCAGGAGTACGGGAGTGGAGTTCACAAAGCTGGCTCTT	1860
DB	1801	TTAAAGGAACCTGAGGGTTCAGGAGTACGGGAGTGGAGTTCACAAAGCTGGCTCTT	1860
QY	1861	GTTTCGATTTATGGGTACTGTTGGAGGAGATTCCAAAGCACCTCCCTCTTTAGTG	1920
DB	1861	GTTTCGATTTATGGGTACTGTTGGAGGAGATTCCAAAGCACCTCCCTCTTTAGTG	1920

1921 GATCTGAGATTCCTTCGAGAGAGTCCCTGGGCCCTGGTGTGGAGAGCGATTGCTA 1980
 1921 GATCTGAGATTCCTTCGAGAGAGTCCCTGGGCCCTGGTGTGGAGAGCGATTGCTA 1980
 1981 CTAGGCCAAGATCTGGAGTTTGAAGAGAAAGAGAGAGGATGAAGTGACGGCCACAAC 2040
 1981 CTAGGCCAAGATCTGGAGTTTGAAGAGAAAGAGAGAGGATGAAGTGACGGCCACAAC 2040
 2041 GACAGCTCANTGGCTTTGAGAGAGACTCTGAGAGAGACTCTCAGGGGGCCAGACCTTGA 2100
 2041 GACAGCTCANTGGCTTTGAGAGAGACTCTGAGAGAGACTCTCAGGGGGCCAGACCTTGA 2100
 2101 CTTCCCTATGGCTCAGTGAACAGAGTCTGGGGGCCGCCGCGCACTAAGTGCGAGAGT 2160
 2101 CTTCCCTATGGCTCAGTGAACAGAGTCTGGGGGCCGCCGCGCACTAAGTGCGAGAGT 2160
 2161 GAAGTTGAGAAACAGACCGAGGTCAGGGGAGGCGAGGGGTGAGAGGCCAGGCCAGCC 2220
 2161 GAAGTTGAGAAACAGACCGAGGTCAGGGGAGGCGAGGGGTGAGAGGCCAGGCCAGCC 2220
 2221 TGTGAGCTGTGTGGGGGCCGACAGTGAAGGGGCCGCTGTGTGGGGCCAGGAGGGCGGGT 2280
 2221 TGTGAGCTGTGTGGGGGCCGACAGTGAAGGGGCCGCTGTGTGGGGCCAGGAGGGCGGGT 2280
 2281 GGGGGGCCCGCTGCCCCACCGTTACTGTTACTCATGCGGCTGTGCGCTTTGCTGTCC 2340
 2281 GGGGGGCCCGCTGCCCCACCGTTACTGTTACTCATGCGGCTGTGCGCTTTGCTGTCC 2340
 2341 CACTACTCAGCACTGAGAGCGGACATGACAGACACAGCGGGGAGAGCGGTTCCG 2400
 2341 CACTACTCAGCACTGAGAGCGGACATGACAGACACAGCGGGGAGAGCGGTTCCG 2400
 2401 TGTGGCGCTGCCCCATPACGNTCAGCCAGTTGTTCACTGACGCGACATACCGGAC 2460
 2401 TGTGGCGCTGCCCCATPACGNTCAGCCAGTTGTTCACTGACGCGACATACCGGAC 2460
 2461 CATACTGGGAGAGCCCTACCGTTGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2520
 2461 CATACTGGGAGAGCCCTACCGTTGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2520
 2521 AACCTGAGGGGATCAGGGGACCCACAGAGGCTTCCCACTCTCTCTCTCTCTCTCTCT 2580
 2521 AACCTGAGGGGATCAGGGGACCCACAGAGGCTTCCCACTCTCTCTCTCTCTCTCTCT 2580
 2581 GGCTTTCGATGCTGCTCAGACCAACCCGCTTCCAGTCCCAAGAGAGAGGGG 2640
 2581 GGCTTTCGATGCTGCTCAGACCAACCCGCTTCCAGTCCCAAGAGAGAGGGG 2640
 2641 ACAATGCCCGAGATCAGAAAAATGCGTGCATCTGCGCAGACTTGAGTCTTCAATGCGCA 2700
 2641 ACAATGCCCGAGATCAGAAAAATGCGTGCATCTGCGCAGACTTGAGTCTTCAATGCGCA 2700
 2701 CCAGGTGGTCCAGTTTCCTGCGCAGACTGCGGAGCTGCGGGGTGAAGGGAGAGCTTG 2760
 2701 CCAGGTGGTCCAGTTTCCTGCGCAGACTGCGGAGCTGCGGGGTGAAGGGAGAGCTTG 2760
 2761 TGTGAACTCGATCCGAACTGCGCAGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2820
 2761 TGTGAACTCGATCCGAACTGCGCAGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2820
 2821 GGACAGAACTGGAGAGGGTGAAGGAGCAGAGCTGGAGCTGCGAGCTGCTGCTGCTGCTG 2880
 2821 GGACAGAACTGGAGAGGGTGAAGGAGCAGAGCTGGAGCTGCGAGCTGCTGCTGCTGCTG 2880
 2881 ATGGAGAGAGGCTGAGAGGGTGGCACTGGGGGACCCAGGGCCCTGCTGAGACAAAGGC 2940
 2881 ATGGAGAGAGGCTGAGAGGGTGGCACTGGGGGACCCAGGGCCCTGCTGAGACAAAGGC 2940
 2941 TTTGCCCTGATTTATGCCCTTTTGGCACTCACTACCCCAACCACTGGCTCGGCAATG 3000
 2941 TTTGCCCTGATTTATGCCCTTTTGGCACTCACTACCCCAACCACTGGCTCGGCAATG 3000

3001 AAGATCTCAGTGGTGAGAAACCTTCCGCTGTGCCGTGTCCATAGCCCTCTGCTCAT 3060
 3001 AAGATCTCAGTGGTGAGAAACCTTCCGCTGTGCCGTGTCCATAGCCCTCTGCTCAT 3060
 3061 CTGGATAAAGCTGAAACCGCAGCGGCTCCACACAGAGGAGAAAGCCCTACAGTGGCCCC 3120
 3061 CTGGATAAAGCTGAAACCGCAGCGGCTCCACACAGAGGAGAAAGCCCTACAGTGGCCCC 3120
 3121 CTTCTGCTCGATGCTGCGTGAACCTTGGCAACCTTGGCAACCTTGGCAACCTTGGCAACCTT 3180
 3121 CTTCTGCTCGATGCTGCGTGAACCTTGGCAACCTTGGCAACCTTGGCAACCTTGGCAACCTT 3180
 3181 GGTGACAAACCTTTCGCTGTAGCTTTCGCACTACAGCTGCAACCTGCAACCTGCAACCTGCA 3240
 3181 GGTGACAAACCTTTCGCTGTAGCTTTCGCACTACAGCTGCAACCTGCAACCTGCAACCTGCA 3240
 3241 AAAAGCTCATATGCTGGGACACACGGGCGGAGAACCTTCCGCTGTGCCACCTCGGCTAT 3300
 3241 AAAAGCTCATATGCTGGGACACACGGGCGGAGAACCTTCCGCTGTGCCACCTCGGCTAT 3300
 3301 ACCACAGGCCACTTGGGACACCTTACAGCTGCAACCTGCAACCTGCAACCTGCAACCTGCA 3360
 3301 ACCACAGGCCACTTGGGACACCTTACAGCTGCAACCTGCAACCTGCAACCTGCAACCTGCA 3360
 3361 GGAGGGCTGCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3420
 3361 GGAGGGCTGCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3420
 3421 TTGAGCACTCGGGGTCCAGAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3480
 3421 TTGAGCACTCGGGGTCCAGAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3480
 3481 TCACCTTGAATTAAGTAACTAGTTTCTTTTACCTGGGCTCTAGGAAATTAGCCCTATGCTCT 3540
 3481 TCACCTTGAATTAAGTAACTAGTTTCTTTTACCTGGGCTCTAGGAAATTAGCCCTATGCTCT 3540
 3541 GCATTTTATACAAATGAACCTAGAAACCACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3600
 3541 GCATTTTATACAAATGAACCTAGAAACCACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3600
 3601 CCACACAGCTAACTAGGCACTATATGACCACTGAAATCCCATGCTGAGGGGGCCAT 3660
 3601 CCACACAGCTAACTAGGCACTATATGACCACTGAAATCCCATGCTGAGGGGGCCAT 3660
 3661 ATAGACCAAGGGAGCTTGTCTTGTAGCTACGTAACAGATGAGCTAAGTATTAGGCGCTTGG 3720
 3661 ATAGACCAAGGGAGCTTGTCTTGTAGCTACGTAACAGATGAGCTAAGTATTAGGCGCTTGG 3720
 3721 ATTACCGCCACTGCTCCAGAGCTATGATGATGATGATGATGATGATGATGATGATGATGAT 3780
 3721 ATTACCGCCACTGCTCCAGAGCTATGATGATGATGATGATGATGATGATGATGATGATGAT 3780
 3781 ACTGTTTTAACTTATTTTCACTGCTTTTATATATAAGGAAACACTAACAGAAAAA 3840
 3781 ACTGTTTTAACTTATTTTCACTGCTTTTATATATAAGGAAACACTAACAGAAAAA 3840
 3841 AAAAAA 3847
 3841 AAAAAA 3847

RESULT 2
 AA235053
 ID AA235053 standard; cDNA; 2651 BP.
 AC XX
 AC AA235053;
 DT 28-FEB-2000 (first entry)
 XX
 DE Mouse CNREB-2L (long isoform) cDNA.
 XX
 KW CNREB-2L; CNRE binding factor; transcription factor; mouse; renin;
 KW negative regulatory element; cardiovascular disease; hypertension;

Db	2461	CCATATAGACCGAGGACCTTGCTTAGCTACGCTACGATGAGCTAAGTGATTAGGGCC	2522
Qy	3717	TTGGATTTCACCGCACTGCTCCAGAGGCTATGGATGAATGGTTGGGAGCTGCCAGCC	3776
Db	2521	TTGGATTTCACCGCACTGCTCCAGAGGCTATGGATGAATGGTTGGGAGCTGCCAGCC	2580
Qy	3777	TTTTACTGTTTTAACTTATTTTTCAGTGCCTTTATTAATAAAGGAACACCTAACAGAAAAAAA	3836
Db	2581	TTTTACTGTTTTAACTTATTTTCAGTGCCTTTATTAATAAAGGAACACCTAACAGAAAAAAA	2640
Qy	3837	AAAAAAAAAAAA 3847	
Db	2641	AAAAAAAAAAAA 2651	
RESULT 3			
AAZ35052			
ID ID AAZ35052 standard; cDNA; 2336 BP.			
XX	AAZ35052;		
XX	28-FEB-2000 (first entry)		
XX	Mouse CNREB-2S (short isoform) cDNA.		
XX	CNREB-2S; CNRE binding factor; transcription factor; mouse; renin;		
KW	negative regulatory element; cardiovascular disease; hypertension;		
KW	congestive heart failure; diagnosis; therapy; ss.		
XX	Mus musculus.		
OS			
XX	Key	Location/Qualifiers	
XX	CD5	674..1978	
XX	FT	/*tag= a	
XX	FT	polya_signal	
XX	FT	2298..2303	
XX	FT	/*tag= b	
XX	XX		
XX	XX	W09955343-A1.	
XX	XX	04-NOV-1999.	
XX	XX	23-APR-1999; 99WO-US008502.	
XX	XX	24-APR-1998; 98US-0082997P.	
XX	XX	(BGHM) BRIGHAM & WOMENS HOSPITAL INC.	
XX	XX	Chen YE, Horiuchi M, Dzau VJ, Tamura K;	
XX	XX	WPI; 2000-062002/05.	
XX	XX	P-PSDE; AAY32376.	
XX	XX	A new regulatory element binding factor used to treat an adverse	
XX	XX	cardiovascular condition.	
XX	XX	Claim 1; Page 83-84; 57pp; English.	
XX	XX	This is the nucleotide sequence of a cDNA clone encoding mouse negative	
XX	XX	regulatory element binding transcription factor 2 (CNREB-2) short isoform	
XX	XX	CNREB-2S. The clone was obtained using yeast one-hybrid cloning and mouse	
XX	XX	kidney cDNA library screening. CNREB-2 binds to the 5' flanking region of	
XX	XX	the renin gene and modulates expression of the gene. 2 Other isoforms	
XX	XX	(see AAY32375 and AAY32377) of CNREB-2 have been identified, the short	
XX	XX	isoform showing strongest binding activity to the CNRE of the renin	
XX	XX	promoter. CNREB-2 shares no significant homology to any other known gene	
XX	XX	or protein. It is a novel zinc finger transcription factor. Expression	
XX	XX	vectors, isolated nucleic acid molecules and CNREB-2 polypeptides are	
XX	XX	claimed. The invention also relates to CNREB-1 polynucleotides (see	
XX	XX	AAZ35040-43), polypeptides (see AAY32374) and inhibitors, and their use	
XX	XX	in claimed methods for decreasing or increasing renin expression, and for	
XX	XX	determining the level of CNREB-1 expression, an increase in expression	
XX	XX	being indicative of an increased susceptibility to development of a renin	
XX	XX	angiotensin system mediated disorder	

XX SQ Sequence 2336 BP; 494 A; 661 C; 737 G; 444 T; 0 U; 0 Other;

Query Match 51.7%; Score 1988.4; DB 3; Length 2336;
 Best Local Similarity 87.8%; Pred. No. 0;
 Matches 2328; Conservative 0; Mismatches 7; Indels 316; Gaps 2;

1198 ATGAGGGGGGAGGGTACACCTTCTGGTTGGCTTGGTTTACCGTCCAGCTTCTG 1257
 1 ATGAGGGGGGAGGGTACACCTTCTGGTTGGCTTGGTTTACCGTCCAGCTTCTG 60

1258 ATTGGTTGGCTAGAGGCGCATCACATCACTTCTGGTCATGCTGGACTGCGCTCCGGT 1317
 61 ATTGGTTGGCTAGAGGCGCATCACATCACTTCTGGTCATGCTGGACTGCGCTCCGGT 120

1318 TGTAAGGCCCATGAATACGATCCCGTGTGCTCTGGTGGCGGAAGCGGAAGCGGGT 1377
 121 TGTAAGGCCCATGAATACGATCCCGTGTGCTCTGGTGGCGGAAGCGGAAGCGGGT 180

1378 ACGAGGTACACAGCTGCTTCTCGAGGGGGGTAGGGGCTCCATGATGGAAGCGGGCG 1437
 181 ACGAGGTACACAGCTGCTTCTCGAGGGGGGTAGGGGCTCCATGATGGAAGCGGGCG 240

1438 GCGCGGGGAGCGACTCAGCTGGATTCCGGGGCGGGG - CAGGGGCTGCCAGGGCGCG 1496
 241 GCGCGGGGAGCGACTCAGCTGGCTCCGGGCGGGGCGGGGCTGCCAGGGGCGG 300

1497 CACCGTGTATGGGGCGGTTCTGGATCCTTAAGAGCAAGAGCCGAGCGGCGCAACT 1556
 301 CACCGTGTATGGGGCGGTTCTGGATCCTTAAGAGCAAGAGCCGAGCGGCGCAACT 360

1557 GGGAGGGCGGGGCGGGAGCGCTCGGCGGGAGCGGCTGGGCTGGAGCGGTCACGA 1616
 361 GGGAGGGCGGGGCGGGAGCGCTCGGCGGGAGCGGCTGGGCTGGAGCGGTCACGA 420

1617 TGCCCCGGAAGCAAGCAAGCCATCCAGCCCGTGAAATGCGAGGGGTCAAAGGTGAGG 1676
 421 TGCCCCGGAAGCAAGCAAGCCATCCAGCCCGTGAAATGCGAGGGGTCAAAGGTGAGG 480

1677 GGTACAGGGGCTTGAAGCGGGAGCAAGGGGTGGGGTCAGTAGAGTGGGTCAGGTCAG 1736
 481 GGTACAGGGGCTTGAAGCGGGAGCAAGGGGTGGGGTCAGTAGAGTGGGTCAGGTCAG 540

1737 GGTACAGGGGACTCTCTCAGGGTTAGGGGCGGATGATCTGGGATCTTCCCTTACCAG 1796
 541 GGTACAGGGGACTCTCTCAGGGTTAGGGGCGGATGATCTGGGATCTTCCCTTACCAG 600

1797 AGTATTAAAGAAACCTGAGGGTCTAGAGTACGGGAAGTGCAGTTCAACAGCTGGCTC 1856
 601 AGTATTAAAGAAACCTGAGGGTCTAGAGTACGGGAAGTGCAGTTCAACAGCTGGCTC 660

1857 CTTGGTTCGGATTATGGGTACTGCTTGGAGGGAGATTCCCAAGCACCTCCCTCTTT 1916
 661 CTTGGTTCGGATTATGGGTACTGCTTGGAGGGAGATTCCCAAGCACCTCCCTCTTT 720

1917 AGTGATACCTGAAGATTCCTTCGAAGAGTCTCGGGGCGGCTGGTGGAGAGCGATT 1976
 721 AGTGATACCTGAAGATTCCTTCGAAGAGTCTCGGGGCGGCTGGTGGAGAGCGATT 780

1977 GCTATAGCCAGATCTGGATTGAGAGAGAGAGAGAGAGATGAGTGAAGCGCA 2036
 781 GCTATAGCCAGATCTGGATTGAGAGAGAGAGAGAGAGATGAGTGAAGCGCA 840

2037 CAACGACCAAGCTCATGGGCTTTGAGAGAGACTCTGAAGAGAGACTCTCAGGGGGCGAGCC 2096
 841 CAACGACCAAGCTCATGGGCTTTGAGAGAGACTCTGAAGAGAGACTCTCAGGGGGCGAGCC 900

2097 TGGACTTCCCTATGGGCTGAGTGAAGAGAGTCTGGGGCGGCGGCGCACTAAGTCCGGA 2156
 901 TGGACTTCCCTATGGGCTGAGTGAAGAGAGTCTGGGGCGGCGGCGCACTAAGTCCGGA 960

2157 GAGTGAAGTGAAGAACCAAGCGGGGTCAGGGAGGCGGAGGGGTGAGAGGCGGCGCC 2216

961 GAGTGAAGTTGAGGAACACGAGCCAGGGGTCAGGGGAGGCCAGGGGTGAGAGGCCAGGCC 1020
 2217 AGCCTGTGAGCTGTGTGGGGGGCCACAGGTGAGGGGCGGCTGTGTGGGGAGGAGGCG 2276
 1021 AGCCTGTGAGCTGTGTGGGGGGCCACAGGTGAGGGGCGGCTGTGTGGGGAGGAGGCG 1080
 2277 GGGTGGGGGGCCCCCGCTGCGCCCAACGGTTACTGTAATCATGCGGGCTGTGGCTTTGCT 2336
 1081 GGGTGGGGGGCCCCCGCTGCGCCCAACGGTTACTGTAATCATGCGGGCTGTGGCTTTGCT 1140
 2337 GTCCCACTACTCGAGCCACCTGAAGCGGCACATGAGAGACACACAGCGGGGAGAACCGTT 2396
 1141 GTCCCACTACTCGAGCCACCTGAAGCGGCACATGAGAGACACACAGCGGGGAGAACCGTT 1200
 2397 CCGCTGTGCGCTGCCCATACGCTNACGCGCAGTTCGTAACCTGAGCGGACATACCG 2456
 1201 CCGCTGTGCGCTGCCCATACGCTNACGCGCAGTTCGTAACCTGAGCGGACATACCG 1260
 2457 CACCCATCTGCGGAGAGCCCTACGTTGTCGCCCACTGCCCTTTGCTGCGAGCAGCT 2516
 1261 CACCCATCTGCGGAGAGCCCTACGTTGTCGCCCACTGCCCTTTGCTGCGAGCAGCT 1320
 2517 GGGCAACCTGAGCGGCGCATCAGCGCACCCACACAGGGGCTCCCACTCTCTGCCAAC 2576
 1321 GGGCAACCTGAGCGGCGCATCAGCGCACCCACACAGGGGCTCCCACTCTCTGCCAAC 1380
 2577 CTGTGCTTTCGATGCTGTGCTCCAGCACACCGGCTCCCACTCCACAGAGCAGGA 2636
 1381 CTGTGCTTTCGATGCTGTGCTCCAGCACACCGGCTCCCACTCCACAGAGCAGGA 1440
 2637 GGGCAACATGCCCCAGCATCAGAAATCGCTGATCTCTGCCAGACTTTGAGTCTTCATGT 2696
 1441 GGGCAACATGCCCCAGCATCAGAAATCGCTGATCTCTGCCAG - - - - - 1484
 2697 GCCACAGGTGTGCCAGTTTCTGCCAGACTGTGGGAGCTGTGGGGTGAAGGGAGAG 2756
 1485 - - - - - 1484
 2757 CTTGTGTGGAATCGGATCCGAACCACTGCCAGAGCTACTGTTCCCTTGGACCTGCGGGG 2816
 1485 - - - - - 1484
 2817 CTGTGACAGGAACCTGGAGGAGGGTGAAGGAGAGGCTGGGAGCTGGCCATGTGTGGGG 2876
 1485 - - - - - 1484
 2877 CTGCATCGAGGAGAGGCTGGAGGGGTGCCACTGGGGGACCCAGGGGCCCTGGTGACAA 2936
 1485 - - - - - 1484
 2937 AGGCTTTGCTGTAGTTTATGCCCCCTTTGCCACTACTACCCCAACCACTGGCTCGGCA 2996
 1485 - - - - - 1485
 2997 CATGAGACTCACAGTGTGAGGAACCTTCCGCTGTGCGCGCTGCCATACGCTCTGC 3056
 1486 CATGAGACTCACAGTGTGAGGAACCTTCCGCTGTGCGCGCTGCCATACGCTCTGC 1545
 3057 TCATCTGGATAACCTGAAACGGCACAGCGGCTCCACACAGGAGAAAGCCCTTACAAAGTG 3116
 1546 TCATCTGGATAACCTGAAACGGCACAGCGGCTCCACACAGGAGAAAGCCCTTACAAAGTG 1605
 3117 CCGCTCTGTCCGTATGCTGTGGCAACCTGGCCACCTCAAGCGTCACTGGTGCATCCA 3176
 1606 CCGCTCTGTCCGTATGCTGTGGCAACCTGGCCACCTCAAGCGTCACTGGTGCATCCA 1665
 3177 CTCTGGTGAACAACTTTTTCGGTGTAGCTTTGCAACTACAGCTGCAACAGAGTATGAA 3236
 1666 CTCTGGTGAACAACTTTTTCGGTGTAGCTTTGCAACTACAGCTGCAACAGAGTATGAA 1725
 3237 CCTCAACAGTCACTATGCTGCAACACGGGCGAGAAAGCCCTTCCGCTGTGCGCAGCTGCC 3296
 1726 CCTCAACAGTCACTATGCTGCAACACGGGCGAGAAAGCCCTTCCGCTGTGCGCAGCTGCC 1785

QY	3297	CTATATCCACAGGCCACTGGGACAACTACAAGCGCTCATCAGAAGGTGCATGGCCATGGTGG	3356
Db	1786	CTATATCCACAGGCCACTGGGACAACTACAAGCGCTCATCAGAAGGTGCATGGCCATGGTGG	1845
QY	3357	AGCAGAGGGCCCTGGTCTCTCTCCCTCGTAGGGCTGGGCCCCACCTCATAGCCACCCCTC	3416
Db	1846	AGCAGAGGGCCCTGGTCTCTCTCCCTCGTAGGGCTGGGCCCCACCTCATAGCCACCCCTC	1905
QY	3417	TGTTTTAGACATCGGGGTCACAGAGCCCTGGGTGCTACTGGTAGCAGGCTCTTCAATTC	3476
Db	1906	TGTTTTAGACATCGGGGTCACAGAGCCCTGGGTGCTACTGGTAGCAGGCTCTTCAATTC	1965
QY	3477	AGACTCACCTTGAACTAACTAGTGTCTTTTACCTGGGGCTCTAGGAATTAGCCCTATGCC	3536
Db	1966	AGACTCACCTTGAACTAACTAGTGTCTTTTACCTGGGGCTCTAGGAATTAGCCCTATGCC	2025
QY	3537	TCCTGCATTTTATACAAATGAATAGAAACCACTTTCCCTTTCTCCCCGTGGTCAAG	3596
Db	2026	TCCTGCATTTTATACAAATGAATAGAAACCACTTTCCCTTTCTCCCCGTGGTCAAG	2085
QY	3597	GGCTCCACACAGACTAACCTTAGCAGCTATATGGACAGCCTGAATCCCATGTGCAGGGG	3656
Db	2086	GGCTCCACACAGACTAACCTTAGGAGCTATATGAGCAGCCTGAATCCCATGTGCAGGGG	2145
QY	3657	CCATATAGACAGGGGACTTGTCTTAGCTCACGTACCAGATGAGCTAAGTGATTAGGCC	3716
Db	2146	CCATATAGACAGGGGACTTGTCTTAGCTCACGTACCAGATGAGCTAAGTGATTAGGCC	2205
QY	3717	TTGGATTACCGGCACGTCTCCACAGAGCTATGGATGAACCTGGTTGGGAGCTGCCAGCC	3776
Db	2206	TTGGATTACCGGCACGTCTCCACAGAGCTATGGATGAACCTGGTTGGGAGCTGCCAGCC	2265
QY	3777	TTTTACTGTTTTAACTTATTTTCAGTGTCTTTATATAAAGGAAACACTAAACAGAAAAAAA	3836
Db	2266	TTTTACTGTTTTAACTTATTTTCAGTGTCTTTATATAAAGGAAACACTAAACAGAAAAAAA	2325
QY	3837	AAAAAAAAAAAA 3847	
Db	2326	AAAAAAAAAAAA 2336	

RESULT 4

ACC83475	ACC83475 standard; cDNA; 1696 BP.
ID	ACC83475 standard; cDNA; 1696 BP.
XX	
XX	ACC83475;
XX	
XX	08-SEP-2003 (first entry)
DT	
DT	
DE	Mouse chondrocyte-derived zinc finger protein Czf-1 partial cDNA.
XX	
KW	Mouse; Czf-1; chondrocyte; zinc finger; osteopathic; antiarthritic;
KW	anti rheumatic; gene; ss.
XX	
OS	Mus sp.
XX	
Key	Location/Qualifiers
FD	3..1355
CDS	/*tag= a
FT	/partial
FT	/product= "Mouse Czf-1"
FT	/note= "No start codon"
FT	

WO200304159-A2.

30-MAY-2003.

20-NOV-2002: 2002WO-IL000925.

20-NOV-2001: 2001US-0331626P.

(PROC-) PROCHON BIOTECH LTD.

XX	Yayon A, Blumenstein S, Harari D;
PI	WPI; 2003-457599/43.
XX	P-P8DB; ABR42914.
DR	
DR	
XX	New chondrocyte-derived zinc finger polypeptides and encoding
PT	polynucleotides, useful for detecting, diagnosing and treating Czf-1
PT	protein-related diseases, such as osteoarthritis and rheumatoid
PT	arthritis.
PT	
XX	Example 1; Fig 4; 64pp; English.
PS	
XX	The present sequence is that of a contig containing most of the murine
CC	Czf-1 coding sequence, lacking only the 5' portion which codes for the N-
CC	terminus of Czf-1 protein. The contig was obtained by screening mouse
CC	expressed sequence tag sequences using rat Czf-1 protein as a sequence
CC	probe. Czf-1 is expressed in a variety of tissue types, particularly
CC	osteoblasts and chondrocytes. It is expressed in the nuclei of
CC	hypertrophic chondrocytes of the growth plate, and exhibits an alternate
CC	expression and intracellular localisation pattern in cells expressing
CC	high, i.e. constitutively active, levels of fibroblast growth factor
CC	receptor (FGFR). It is up-regulated in osteoarthritic cartilage tissue
CC	and serves as a marker for osteoarthritis. Czf-1 has been found in rats,
CC	mice and humans. Sequence identity for nucleotide and amino acid
CC	sequences is above 85%. Czf-1 polynucleotides, polypeptides and
CC	antibodies can be used in the characterisation, diagnosis and treatment
CC	of FGFR-related and skeletal diseases and disorders, such as
CC	osteoarthritis, rheumatoid arthritis, and cartilage related diseases
CC	
XX	Sequence 1696 BP; 346 A; 529 C; 487 G; 334 T; 0 U; 0 Other;
SQ	
	Query Match 43.5%; Score 1671.8; DB 7; Length 1696;
	Best Local Similarity 99.5%; Pred. No. 0;
	Matches 1687; Conservative 0; Mismatches 8; Indels 1; Gaps 1;
QY	2135 GCGCCGCGCACAATACTGGAGAGTGAAGTTGAGAAACAGGCAGGGGTCCAGGGAGG 2194
Db	1 GCGCCGCGCACTAAGTCGGAGAGTGAAGTTGAGAAACAGGCAGGGGTCCAGGGAGG 60
QY	2195 CCAGGGGTGAGAGGCCAGGCCCACTGTCAGCTGTGTGGGGGGCCGACAGGTGAGGGGC 2254
Db	61 CCAGGGGTGAGAGGCCAGGCCCACTGTCAGCTGTGTGGGGGGCCGACAGGTGAGGGGC 120
QY	2255 CGTGTTGTGGGCAGAGGGCGGGGTGGGGGGCCCCCGCTGCCCCCACAGGTTACTGTACT 2314
Db	121 CGTGTTGTGGGCAGAGGGCGGGGTGGGGGGCCCCCGCTGCCCCCACAGGCTACTGTACT 180
QY	2315 CATCGCGGTGTGCGCTTTTCGTGCCCATCTACTCAGCCACCTGAGCGGCACATGCAGA 2374
Db	181 CATGCGCGGTGTGCGGCTTTCGTGCCCATCTACTCAGCCACCTGAGCGGCACATGCAGA 240
QY	2375 CACACAGCGGGAGAGACCGTTCGGCTGTGGCCGCTGCCCATACGNTACGCCAGTTCG 2434
Db	241 CACACAGCGGGAGAGACCGTTCGGCTGTGGCCGCTGCCCATACGCTCAGCCAGCTCG 300
QY	2435 TCAACTCAGCGGACATACCAGCACCGTACTGGCGAGAGCCCTACCGTTGCCCACT 2494
Db	301 TCAACTCAGCGGACATACCAGCACCGTACTGGCGAGAGCCCTACCGTTGCCCACT 360
QY	2495 GCCCCTTTGCTGCAGCAGCTGGGAAACTCTGAGGCGGCTCAGCGCACCCACACAGGGC 2554
Db	361 GCCCCTTTGCTGCAGCAGCTGGGAAACTCTGAGGCGGCTCAGCGCACCCACACAGGGC 420
QY	2555 CTCCTCATCTCCCTGCCAATCTGTGGCTTTTGATGCTGTGCTTCCAGACCAACCCGGC 2614
Db	421 CTCCTCATCTCCCTGCCAATCTGTGGCTTTTGATGCTGTGCTTCCAGACCAACCCGGC 480
QY	2615 CTCCTCAGTCCACAGAGCAGAGGGGACAAATGCCCGACGATCAGAAAATGCGCTGATCC 2674
Db	481 CTCCTCAGTCCACAGAGCAGAGGGGACAAATGCCCGACGATCAGAAAATGCGCTGATCC 540
QY	2675 TGCCAGACTTTGATTCTTCATGTGCCACAGGTTGGTGCCAGTTTTCTTGCCAGACTGTGGC 2734

Db 541 TGCCAGACTTGAGTCTTCATGTGCGCACAGGTGTGCCAGTTTCTGCCAGACTGTGGGC 600
Qy 2735 AGCTGCGGGGTGAAGGGGAGAGCTTGTGTGAACTGGATCCGAACCACTGCCAGAGCTAC 2794
Db 601 AGCTGCGGGGTGAAGGGGAGAGCTTGTGTGAACTGGATCCGAACCACTGCCAGAGCTAC 660
Qy 2795 TGTTCCTTGGACCTGCGGGGCTGTGGACGGAACCTGGAGAGGGTGTAGGGGAGCAGGC 2854
Db 661 TGTTCCTTGGACCTGCGGGGCTGTGGACGGAACCTGGAGAGGGTGTAGGGGAGCAGGC 720
Qy 2855 TGGAGCTGCCATGTGTGGGCGCTGCATGCGAGAGAGGCTGGAGGGGTGCCACTGGGG 2914
Db 721 TGGAGCTGCCATGTGTGGGCGCTGCATGCGAGAGAGGCTGGAGGGGTGCCACTGGGG 780
Qy 2915 GA-CCCCAGGCGCTGTGACAAAGGCTTTCCTGTGTATGCCCCCTTGGCACTCAC 2973
Db 781 GACCCCCAGGCGCTGTGACAAAGGCTTTCCTGTGTATGCCCCCTTGGCACTCAC 840
Qy 2974 TACCCCAACCACTGGCTGCGGCACATGAAGACTCACAGTGTGAGAAACCTTCCGCTGT 3033
Db 841 TACCCCAACCACTGGCTGCGGCACATGAAGACTCACAGTGTGAGAAACCTTCCGCTGT 900
Qy 3034 GCGCGCTGCATACGCTCTGCTCATCTGATTAACCTGAAACGCGACCGAGCGGTCCAC 3093
Db 901 GCGCGCTGCATACGCTCTGCTCATCTGATTAACCTGAAACGCGACCGAGCGGTCCAC 960
Qy 3094 ACAGGAGAAAGCCCTACAGTGGCCCTCTGCTCGGTATGCTGTGGCAACCTGGCCAAAC 3153
Db 961 ACAGGAGAAAGCCCTACAGTGGCCCTCTGCTCGGTATGCTGTGGCAACCTGGCCAAAC 1020
Qy 3154 CTCAGGCTCATGTGCGCATCCATCTGTGTGACAAACCTTTTGGTGTAGCCCTTGGCAAC 3213
Db 1021 CTCAGGCTCATGTGCGCATCCATCTGTGTGACAAACCTTTTGGTGTAGCCCTTGGCAAC 1080
Qy 3214 TACAGCTGCACACAGATGATGAACCTCAAGCTCATATGCTGCACACACGGGCGAGAG 3273
Db 1081 TACAGCTGCACACAGATGATGAACCTCAAGCTCATATGCTGCACACACGGGCGAGAG 1140
Qy 3274 CCGTTCCGCTGTGCCACCTGCGCCTATACACAGGCCACTGGGACAACTACAAGCGTCA 3333
Db 1141 CCGTTCCGCTGTGCCACCTGCGCCTATACACAGGCCACTGGGACAACTACAAGCGTCA 1200
Qy 3334 CAGAAGTGCATGCGCATGTGTGAGCAGAGGGGCTGTCTCTGCCCCCTGAGGGGTGG 3393
Db 1201 CAGAAGTGCATGCGCATGTGTGAGCAGAGGGGCTGTCTCTGCCCCCTGAGGGGTGG 1260
Qy 3394 GCCCCACCTCATAGCCACCCCTCTGTTTGTAGCACTCGGGTCCAGCAGCCCTGGGTGCT 3453
Db 1261 GCCCCACCTCATAGCCACCCCTCTGTTTGTAGCACTCGGGTCCAGCAGCCCTGGGTGCT 1320
Qy 3454 ACTGTAGCAGGGCTTTTCATTCAGACTCACCTTGAACCTAGTGGTCTTTTACCTGGG 3513
Db 1321 ACTGTAGCAGGGCTTTTCATTCAGACTCACCTTGAACCTAGTGGTCTTTTACCTGGG 1380
Qy 3514 GCTCTAGAAATTAGCCCTATGCCCTCTGCTATATACAAATGAACCTAGAAACCACTTT 3573
Db 1381 GCTCTAGAAATTAGCCCTATGCCCTCTGCTATATACAAATGAACCTAGAAACCACTTT 1440
Qy 3574 CCGTTTCCCGCGTGTGAGGGGCTCCACACAGACTAAGCTAGGCACTATATGACCA 3633
Db 1441 CCGTTTCCCGCGTGTGAGGGGCTCCACACAGACTAAGCTAGGCACTATATGACCA 1500
Qy 3634 GCCTGAATCCCATGTCAGGGGGCCATATACACAGGGGACTTCTTAGTCACTACGTACC 3693
Db 1501 GCCTGAATCCCATGTCAGGGGGCCATATAGACAGGGGACTTCTTAGTCACTACGTACC 1560
Qy 3694 AGATGAGCTAAGTATAGGCGCTTGGATTACCGCCACTGCTCCAGAGGCTATGATG 3753
Db 1561 AGATGAGCTAAGTATAGGCGCTTGGATTACCGCCACTGCTCCAGAGGCTATGATG 1620
Qy 3754 AACTGTTGGGAGCTGCCAGCCCTTACTGTTTAACTATTTCAGTGGCTTTATATATA 3813

Db 1621 AACTGTTGGGAGCTGCCAGCCTTTACTGTTTAACTATTTCAGTGGCTTTATATA 1680
Qy 3814 AGGAAACACTAACAGA 3829
Db 1681 AGGAAACACTAACAGA 1696
RESULT 5
ACC83474
ID ACC83474 standard; cDNA; 1860 BP.
XX
AC ACC83474;
XX
DT 08-SEP-2003 (first entry)
XX
DE Rat chondrocyte-derived zinc finger protein Czf-1 cDNA.
XX
KW Rat; Czf-1; chondrocyte; zinc finger; osteopathic; antiarthritic;
KW antirheumatic; gene; ss.
XX
OS Rattus sp.
XX
Key Location/Qualifiers
CDS 176..1801
/*tag= a
FT /product= "Rat Czf-1"
XX
PN W02003044159-A2.
XX
PD 30-MAY-2003.
XX
PF 20-NOV-2002; 2002WO-IL000925.
XX
PR 20-NOV-2001; 2001US-0331626P.
XX
(PROC-) PROCHON BIOTECH LTD.
XX
PI Yaron A, Blumenstein S, Harari D;
XX
DR WPI; 2003-457599/43.
DR P-PSDB; ABR42913.
XX
PT New chondrocyte-derived zinc finger polypeptides and encoding
PT polynucleotides, useful for detecting, diagnosing and treating Czf-1
PT protein-related diseases, such as osteoarthritis and rheumatoid
PT arthritis.
XX
Example 1; Fig 1; 64pp; English.
XX
CC The present sequence is that of cDNA encoding a novel rat zinc finger
CC protein, designated Czf-1. The cDNA was isolated from a glioma cDNA
CC library using a probe from the 3' end of the Czf-1 gene. Czf-1 is
CC expressed in a variety of tissue types, particularly osteoblasts and
CC chondrocytes. It is expressed in the nuclei of hypertrophic chondrocytes
CC of the growth plate, and exhibits an alternate expression and
CC intracellular localization pattern in cells expressing high, i.e.
CC constitutively active, levels of fibroblast growth factor receptor
CC (FGFR). It is up-regulated in osteoarthritic cartilage tissue and serves
CC as a marker for osteoarthritis. Czf-1 has been found in rats, mice and
CC humans. Sequence identity for nucleotide and amino acid sequences is
CC above 85%. Czf-1 polynucleotides, polypeptides and antibodies can be used
CC in the characterization, diagnosis and treatment of FGFR-related and
CC skeletal diseases and disorders, such as osteoarthritis, rheumatoid
CC arthritis, and cartilage related diseases
XX
SQ Sequence 1860 BP; 365 A; 564 C; 601 G; 329 T; 0 U; 0 Other;
Query Match 39.1%; Score 1504.4; DB 7; Length 1860;
Best Local Similarity 95.4%; Pred. No. 0;
Matches 1560; Conservative 0; Mismatches 72; Indels 3; Gaps 1;
Qy 1917 AGTGATACCTGAGATTCCTTCAGCAAGGTCCTCGGCGCCCTGGGTGGAGAGCATTT 1976

Db 229 AGTGGATCTGAAGATTCTCTTCGACGAAGGACCTGGGCTCTGGTGTGGAGAGTGATT 288
Qy 1977 GCTACTAGCCCAAGATCTGGAGTTTAAAGAGAAAGGAAGGATGAAGTGACGGCA 2036
Db 289 GCTACTAGCCCAAGATCTGGAGTTTAAAGAGAAAGGAAGGATGAAGTGACGGCA 348
Qy 2037 CAACGACCAAGCTCATGGGCTTTGAGAGAGACTCTCAAGAGAGCTCTCAGGGGGCCAGACC 2096
Db 349 CAACGACCAAGCTCATGGGCTTTGAGAGAGACTCTCAAGAGAGCTCTCAGGGAGCCAGACC 408
Qy 2097 TGGAGCTTCCCTATGGGCTGAGTGACGACGAGTCTGGGGGCGGCCGCGACATAAGTGCGA 2156
Db 409 TGGAGCTTCCCTATGGGCTGAGTGACGATGAGTCTGGGGGCGGCCGCGACATAAGTGCGA 468
Qy 2157 GAGTGAAGTTGAGAACACAGCCAGGGGTCAGGGAGGCCAGGGGTGAGAGGCCAGGGCC 2216
Db 469 AAGTGAAGTTGAGAACCCAGCCAGGGGTCCAGGGAGGCCAGGGGTGAGAGGCCAGGGCC 528
Qy 2217 AGCCTGTGAGCTGTGTGGGGGCGCCACAGGTGAGGGGCGCGTCTGTGGGGGCGAGAGGGCG 2276
Db 529 AGCCTGTGAGCTGTGTGGGGGCGCCACAGGTGAGGGGCGCGTCTGTGGGGGCGAGAGGGCG 588
Qy 2277 GGGTGGGGGCGCCCGCTGCCCGCCACGCTTACTGTAATCTGCGGCTGTGGCTTTGCT 2336
Db 589 GGGTGGGGGCGCCCGCTGCCCGCCACGCTTACTGTAATCTGCGGCTGTGGCTTTGCT 648
Qy 2337 GTCCCACTACTGAGCCACTGAAGCGGCACATGACAGACACAGCGGGGAGAACCGTT 2396
Db 649 GTCCCACTACTGAGCCACTGAAGCGGCACATGACAGACACAGCGGGGAGAACCGTT 708
Qy 2397 CGCTGTGGCGGTGCCCATAGCNCVTCAGCCAGTTCGTCAACCTGACCGCATACCGG 2456
Db 709 CGCTGTGGCGGTGCCCATAGCNCVTCAGCCAGTTCGTCAACCTGACCGCATACCGG 768
Qy 2457 CACCCATCTGGCGAGAGCCCTACCGTTGTCCTCCACTGCCCTTTGCTGTCAGAGCCT 2516
Db 769 CACCCATCTGGCGAGAGCCCTACCGTTGTCCTCCACTGCCCTTTGCTGTCAGAGCCT 828
Qy 2517 GGGCAACCTGAGCGGCATCAGCGCACCCACAGGGCTCCCACTCTCTCCCTGCCCAAC 2576
Db 829 GGGCAACCTGAGCGGCATCAGCGCACCCACAGGGCTCCCACTCTCTCCCTGCCCAAC 888
Qy 2577 CTGTGCTTCGATGTGTGCTCAGACCAACCGGCTCCAGTCCCAAGCCAGAGCGAGA 2636
Db 889 CTGTGCTTCGATGTGTGCTCAGACCAACCGGCTCCAGTCCCAAGCCAGAGCGAGA 948
Qy 2637 GGGGACAAATCCCGCCAGCATCAGAAATGCGCTGATCTCTGCGAGACTTGAGTCTTCATGT 2696
Db 949 AGGACAAATCCCGCCAGCATCAGAAATGCTCTCATCTCCCGAGACTTGAGTCTTCATGT 1008
Qy 2697 GCACCAAGTGTGTCAGTTTCTGCGAGACTGTGGGAGCTCGGGGTGAAGGGAGAG 2756
Db 1009 GCCACCAAGTGTGTCAGTTTCTGCGAGACTTGGGGAGCTCGGGGTGAAGGGAGAG 1068
Qy 2757 CTTGTGTGGAACCTGGATCCGAACACTGCGAGAGTACTGTTCCCTTGGACCTGCCGGGG 2816
Db 1069 TTTGTGTGGAACCTGGATCAGAACCACTGCGAGAGTACTGTTCCCTTGGACCTGCCGGGG 1128
Qy 2817 CTGTGGAAGAACTGGAGAGGGTGAAGCCAGAGCTGGGAGCTGCCATGTGTGGCG 2876
Db 1129 CTGTGGAAGAACTGGAGAGGGTGAAGCCAGAGCTGGGAGCTGCCATGTGTGGCG 1188
Qy 2877 CTGATCGGAGAGAGCTGGAGGGTGGCACTGGGGGACCCAGGGCCCTGTCACAA 2936
Db 1189 CTGATCGGAGAGAGCTGGAGGGTGGCACTGGGGGACCCAGGGCCCTGTCACAA 1248
Qy 2937 AGGCTTTGCTGTAGTTTATGCCCCCTTTGCCACTCACTACCCCAACACCTGCTCGGA 2996
Db 1249 AGGCTTTGCTGTAGTTTATGCCCCCTTTGCTACTCACTACCCCAACACCTGCTCGGA 1308
Qy 2997 CATGAAGCTCAGAGTGGTGAAGAACCTTTCCGCTGTGCGGCTGTCCATGCGCTCTGC 3056
Db 1309 CATGAAGCTCAGAGTGGTGAAGAACCTTTCCGCTGTGCGGCTGTCCATGCGCTCTGC 1368

Qy 3057 TCATCTGGATAACCTGAAGCGCACAGCGCGTCCACAGGAGAAAAGCCCTACAAGTG 3116
Db 1369 GCATCTGGATAACCTGAAGCGCACAGCGCGTCCACAGGAGAAAAGCCCTACAAGTG 1428
Qy 3117 CCCTCTGTCCGTATGCTGTGGCAACCTGTGGCAACCTGTGGCAACCTCAAGCGTCAAGTGTATCCA 3176
Db 1429 CCCTCTGTCCGTATGCTGTGGCAACCTGTGGCAACCTGTGGCAACCTCAAGCGTCAAGTGTATCCA 1488
Qy 3177 CTCTGGTGACAAACCTTTTCGGTGTAGCCTTTGCAACTACAGCTGCAACAGAGATGAA 3236
Db 1489 CTCTGGTGACAAACCTTTTCGGTGTAGCCTTTGCAACTACAGCTGCAACAGAGATGAA 1548
Qy 3237 CCTCAACGCTCATATGCTGCGACACACGCGCGGAGAAAGCCCTTCGGTGTGCCACCTGCGC 3296
Db 1549 CCTCAACGCTCATATGCTGCGACACACGCGCGGAGAAAGCCCTTCGGTGTGCCACCTGTCG 1608
Qy 3297 CTATACACAGGCGCATGTGGGACAACTACAAGCGTCAACAAGGTGATGCGCATGGTGG 3356
Db 1609 TTACACACAGGCGCATGTGGGACAACTACAAGCGTCAACAAGGTGATGCGCATGGTGG 1668
Qy 3357 AGCAGGAGGCGCTGCTCTCTGCCCCCTGAGGGCTGGGCCCCACCTCATAGCCACCCCTC 3416
Db 1669 AGCAGGAGGCGCTGCTCTCTGCCCCCTGAGGGCTGGGCCCCACCTCATAGCCACCCCTC 1728
Qy 3417 TGTCTTGAAGCACTGGGCTCCAGAGCCCTGGGTGCTACTGTGTAGAGGGCTCTTCAATTC 3476
Db 1729 TGTCTTGAAGCACTGGGCTCCAGCGCCCTGGGTGCTACTGTGTAGAGGGCTCTTCAATTC 1788
Qy 3477 AGACTCACTTGAACCTAAGTGTCTTTTACCTGGGGCTTAGGAATTAGCCCTATGCC 3536
Db 1789 AGACTCACTTGAACCTAAGTGTCTTTTACCTGGGGCTTAGGAATTAGCCCTATGCC 1845
Qy 3537 TCCTGCACTTTTATAC 3551
Db 1846 TCCTGCACTTTTATAC 1860

RESULT 6
ABAB2987
ID ABAB2987 standard; DNA; 1975 BP.
XX
AC ABA82987;
XX
DT 05-FEB-2002 (first entry)
XX
DE Human transcription factor TRFX-14 coding sequence.
XX
KW Human; transcription factor; TRFX; cell proliferative disease;
KW autoimmune disease; inflammation; neurological disease;
KW developmental disorder; cancer; AIDS; infection; cytostatic; anti-HIV;
KW neuroprotective; antiinflammatory; gene therapy; ds.
XX
OS Homo sapiens.
XX
PN WO200172777-A2.
XX
PD 04-OCT-2001.
XX
PF 13-MAR-2001; 2001WO-US008117.
XX
PR 13-MAR-2000; 2000US-0188986P.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Hillman JL, Baughn MR, Yue H, Lal P, Lu DM, Patterson C;
PI Azimzai Y, Bandman O, Tang YT, Machur P, Shah P, Au-Young J;
PI Reddy R;
XX
DR WPI; 2001-570896/64.
XX
DR P-PSDB; AB50163.
XX
PT Novel transcription factor polypeptides, used to treat diseases

PT associated with altered activity and expression of TRFX, and to screen
 PT for agents capable of modulating its activity.
 XX
 PS Claim 11; Page 261; 327pp; English.
 XX The present sequence is the coding sequence for a human transcription
 CC factor. The transcription factor and its coding sequence are useful in
 CC the diagnosis, treatment and prevention of diseases associated with
 CC altered expression of the transcription factor e.g. cell proliferative,
 CC autoimmune/inflammatory, neurological and developmental disorders. A
 CC number of specific disorders/diseases are given in the specification,
 CC including: arteriosclerosis, cirrhosis, hepatitis, cancers, AIDS,
 CC allergies, anemia, asthma, autoimmune thyroiditis, bronchitis, atopic
 CC dermatitis, diabetes mellitus, emphysema, Goodpasture's syndrome, gout,
 CC Grave's disease, multiple sclerosis, osteoarthritis, pancreatitis,
 CC psoriasis, rheumatoid arthritis, systemic lupus erythematosus, ulcerative
 CC colitis, uveitis, Alzheimer's disease, Huntington's disease, Parkinson's
 CC disease, stroke, and viral, bacterial, fungal and protozoal infections
 XX
 SQ Sequence 1975 BP; 396 A; 625 C; 612 G; 342 T; 0 U; 0 Other;
 Query Match 38.9%; Score 1498; DB 5; Length 1975;
 Best Local Similarity 88.6%; Pred. No. 0;
 Matches 1707; Conservative 0; Mismatches 176; Indels 44; Gaps 6;
 QY 1917 AGTGATGATGAGTCTCTGCGAGAGGTCCTGGGCGCCCTGGTGTGGAGCGGATTT 1976
 DB 86 AGTGATGATGAGTCTCTGCGAGAGGTCCTGGGCGCCCTGGTGTGGAGCGGATTT 145
 QY 1977 GCTACTAGCCCAAGATCTGGAGTTTGAAGAGAGAGAGAGAGAGATGAAGTGACGCCCA 2036
 DB 146 GCTACTAGCCCAAGATCTGGAGTTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 205
 QY 2037 CAACACAGATCATGGCTTTGAGAGAGATCTGAGAGAGATCTGAGAGAGATCTGAGAGAG 2096
 DB 206 CAGTACAGATCATGGCTTTGAGAGAGATCTGAGAGAGATCTGAGAGAGATCTGAGAGAG 265
 QY 2097 TGGACTTCCCTATGGGCTGAGTGACGAGAGTCTGGGGGGCGCGGCACTAAAGTGCGGA 2156
 DB 266 TGGGCTTCCCTATGGCTGAGCGAGATGAGTCTGGGGGGCGCGGCACTAAAGTGCGGA 325
 QY 2157 GAGTGAAGTTGAGAAACAGCAGAGGGTTCAGGGAGGCGCAGGGGTGAGAGGCGAGGCC 2216
 DB 326 GAGTGAAGTTGAGAGAGCAGCAGGGGTTCAGGGAGGCGCAGGGGTGAGAGGCGAGGCC 385
 QY 2217 AGCCTGTGAGCTGTGGGGGGCGCAGAGTGTGGGGGGCGCGGCTGTTGTGGGGCAGAGGGCG 2276
 DB 386 AGCCTGCCAGCTGTGGGGGGCGCAGAGTGTGGGGGGCGCGGCTGTTGTGGGGCAGAGGGCG 445
 QY 2277 GGGTGGGGGGCGCGCTGCGCCCGCAGGTGATGATGATGATGATGATGATGATGATGATGAT 2336
 DB 446 GGGTGGGGGGCGCGCTGCGCCCGCAGGTGATGATGATGATGATGATGATGATGATGATGAT 505
 QY 2337 GTCCCACTACTGAGCCCACTGAGCGGCACATGACAGACACAGCGGGGAGAGCGGTT 2396
 DB 506 GTCCCACTACTGAGCGGCACCTGAGCGGCACATGACAGACACAGCGGGGAGAGCGGTT 565
 QY 2397 CGCGTGTGGCGCTGCCCATAGCAGCGGCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTC 2456
 DB 566 CGCGTGTGGCGCTGCCCATAGCAGCGGCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTC 625
 QY 2457 CACCCATAGTGGGAGAGCGCTACCGTTGTCCTCCCACTGGCCCTTGGCTTCAGCAGCGCT 2516
 DB 626 CACCCATAGTGGGAGAGCGCTACCGTTGTCCTCCCACTGGCCCTTGGCTTCAGCAGCGCT 685
 QY 2517 GGGCAACCTGAGCGGCGATCAGCGCACCCACACAGGGGCTTCCCACTTCTCCCTGCCAAC 2576
 DB 686 GGGCAACCTGAGCGGCGATCAGCGCACCCACACAGGGGCTTCCCACTTCTCCCTGCCAAC 745
 QY 2577 CTGTGGCTTTCAGTGTGCTCCAGGACCAACCGGGCTTCCAGTCCCAAGAGGAGGA 2636
 DB 746 CTGTGGCTTTCAGTGTGCTCCAGGACCAACCGGGCTTCCAGTCCCAAGAGGAGGA 805

2637 GGGGCAATGCCCCGACGATCAGAAAAATGCGCTGATCCTGCCAGACTTGAGTCTTCATGT 2696
 DB 806 GGGGGCGGTGCCCCGCGACCTGAAAGATGCTCTGCTCTCCAGATTTGAGCCTCATGT 865
 QY 2697 GCACACAGTGTGCCAGTTCTTCTGCCAGATGCTGGGCGAGCTGCGGGTGAAGGGAGAG 2756
 DB 866 GCCACAGTGTGCCAGTTCTTCTGCCAGATGCTGGGCGAGCTGCGGGTGAAGGGAGAG 925
 QY 2757 CTTGTGTGAACTGGATCCGAACCACTGCCAGAGCTACTGTTCCCTTGGACCTGCCGGG 2816
 DB 926 CTTGTGTGAACTGGATCCGAACCACTGCCAGAGCTGCTATTCCCTTGGACCTGCCGGG 985
 QY 2817 CTGTGACAGAACTGGAGGAGGCTGAGGAGAGGCTGAGGAGAGGCTGGGAGCTGCATGTG 2876
 DB 986 CTGTGACAGAACTGGAGGAGGCTGAGGAGAGGCTGAGGAGAGGCTGGGAGCTGCATGT 1045
 QY 2877 CTGCATGCGCAGAGAGGCTGAGGAGGCTGAGGAGAGGCTGAGGAGAGGCTGGTGAACA 2936
 DB 1046 CTGCATGCGCAGAGAGGCTGAGGAGGCTGAGGAGAGGCTGAGGAGAGGCTGGTGAACA 1105
 QY 2937 AGCTTTGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2996
 DB 1106 AGCTTTGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1165
 QY 2997 CATGAAGACTCACAGTGTGAGAAACCTTCCGCTGTGCGCTGTGCGCTGTGCGCTGTGCG 3056
 DB 1166 CATGAAGACTCACAGTGTGAGAAACCTTCCGCTGTGCGCTGTGCGCTGTGCGCTGTGCG 1225
 QY 3057 TCATCTGGATACCTGAAACCGGCAACAGCGCTTCCACACAGAGAGAAAGCCCTTCAAGTG 3116
 DB 1226 TCATCTGGATACCTGAAACCGGCAACAGCGCTTCCACACAGAGAGAAAGCCCTTCAAGTG 1285
 QY 3117 CCGCTCTGCTGATGCTGTGGCAACCTGCGCAACCTCAAGCTCATGCTGCGATCCA 3176
 DB 1286 CCGCTCTGCTGATGCTGTGGCAACCTGCGCAACCTCAAGCTCATGCTGCGATCCA 1345
 QY 3177 CTCTGTGACAAACCTTTTGGTGTAGCTTGTGCAACTACAGCTGCAACAGAGATGAA 3236
 DB 1346 CTCTGTGACAAACCTTTTGGTGTAGCTTGTGCAACTACAGCTGCAACAGAGATGAA 1405
 QY 3237 CTTCAACCTCATGCTGCGACACAGCGGCGAGAGCCCTTCCGCTGTGCGCTGTGCGCTGCG 3296
 DB 1406 CTTCAACCTCATGCTGCGGCGACACAGCGGCGAGAGCCCTTCCGCTGTGCGCTGTGCGCTGCG 1465
 QY 3297 CTATACACAGGCGCACTGGGACAACTACAAGCTCATCAAGAGGTGATGCGCATGTTGG 3356
 DB 1466 CTATACACAGGCGCACTGGGACAACTACAAGCTCATCAAGAGGTGATGCGCATGTTGG 1525
 QY 3357 AGCAGAGGCGCTGTGCTCTGCGCTGAGGGCTGGGCGCCACCTCATAGCCGCCCTC 3416
 DB 1526 AGCAGAGGCGCTGTGCTCTGCGCTGAGGGCTGGGCGCCACCTCATAGCCGCCCTC 1585
 QY 3417 TGTGTTGAGCACTCGGGGTCCAGAGCCCTGCGGTGCTACTGCTGAGCGGCTCTTCAATTC 3476
 DB 1586 TGTGTTGAGCACTCGGGGTCCAGAGCCCTGCGGTGCTACTGCTGAGCGGCTCTTCAATTC 1645
 QY 3477 AGACTCACTTGAACCTAAGTGTGCTTCTTCTGCGCTGTGAGGGCTGAGGATAGCCCTATGCC 3536
 DB 1646 AGACTCACTTGAACCTAAGTGTGCTTCTTCTGCGCTGTGAGGGCTGAGGATAGCCCTATGCC 1672
 QY 3537 TCCTGCAATTTTACAAATGAACTAGAAACCACTTCTCCCTTCTCCCGGCTGTGTCAGG 3596
 DB 1673 TCCTGCAATTTTACAAATGAACTAGAAACCACTTCTCCCTTCTCCCGGCTGTGTCAGG 1728
 QY 3597 GGTTCACACAGACTAACTAGGCACTATATGAGCAAGCTGAACTGCTGAGGGGG 3656
 DB 1729 GGTTCACACAGACTAACTAGGCACTATATGAGCAAGCTGAACTGCTGAGGGGG 1788
 QY 3657 -CCATATAGACAGGGGA-CTTGTCTTAGCTACGTACAGATGAGCTAGCTAAGTATGAGG 3714
 DB 1789 CCATATGACAGGGGA-CTTGTCTTAGCTACGTACAGATGAGCTAGCTAAGTATGAGG 1848
 QY 3715 CTTTGAATTCACGCCCACTGCTCCAGAGGCTATGAGTGAATGCTGTTGGAG-CTGCCCA 3773

Db 1849 CCTGTATTCACCTCACTGCCCCAGGGCTGTGCACAAACCGCTGGGGAGTGCACA 1908
QY 3774 GCCTTTTAC-----TGTTTTAACCTATTTCAGTCTTTTATATAAAGGAAACACTAACAGA 3829
Db 1909 GCCTCCACCTCTTTTATTAACTTATTTCAGTCTTTTATATAAAGGAAACACTAACAAA 1968
QY 3830 AAAAAAA 3836
Db 1969 AAAAAAA 1975
RESULT 7
ACC83473
ID ACC83473 standard; cDNA; 1855 BP.
XX AC ACC83473;
DT 08-SEP-2003 (first entry)
XX DE Human chondrocyte-derived zinc finger protein Czf-1 cDNA.
XX KW Human; Czf-1; chondrocyte; zinc finger; osteopathic; antiarthritic;
XX KW antirheumatic; chromosome 2p22-23; gene; ss.
XX OS Homo sapiens.
XX PH Key Location/Qualifiers
XX FT CDS 171..1796
XX FT /*tag= a
XX FT /product= "Human Czf-1"
XX PN WO2003044159-A2.
XX PD 30-MAY-2003.
XX PF 20-NOV-2002; 2002WO-IL000925.
XX PR 20-NOV-2001; 2001US-0331626P.
XX PA (PROC-) PROCHON BIOTECH LTD.
XX PI Yayon A, Blumenstein S, Harari D;
XX WPI; 2003-457599/43.
XX DR P-PSDB; ABR42912.
XX PT New chondrocyte-derived zinc finger polypeptides and encoding
XX PT polynucleotides, useful for detecting, diagnosing and treating Czf-1
XX PT protein-related diseases, such as osteoarthritis and rheumatoid
XX PT arthritis.
XX PS Claim 12; Fig 1; 64pp; English.
XX CC The present sequence is that of cDNA encoding a novel human zinc finger
XX CC protein, designated Czf-1. Czf-1 is expressed in a variety of tissue
XX CC types, particularly osteoblasts and chondrocytes. It is expressed in the
XX CC nuclei of hypertrophic chondrocytes of the growth plate, and exhibits an
XX CC alternate expression and intracellular localisation pattern in cells
XX CC expressing high, i.e. constitutively active, levels of fibroblast growth
XX CC factor receptor (FGFR). It is up-regulated in osteoarthritic cartilage
XX CC tissue and serves as a marker for osteoarthritis. Czf-1 has been found in
XX CC rats, mice and humans. Sequence identity for nucleotide and amino acid
XX CC sequences is above 85%. The human gene maps to chromosome 2p22-23. Czf-1
XX CC polynucleotides, polypeptides and antibodies can be used in the
XX CC characterisation, diagnosis and treatment of FGFR-related and skeletal
XX CC diseases and disorders, such as osteoarthritis, rheumatoid arthritis, and
XX CC cartilage related diseases
XX SQ Sequence 1855 BP; 344 A; 590 C; 619 G; 302 T; 0 U; 0 Other;
Query March 36.2%; Score 1392; DB 7; Length 1855;
Best Local Similarity 92.6%; Pred. No. 0;

Matches 1461; Conservative 0; Mismatches 116; Indels 0; Gaps 0;
QY 1917 AGTGGATACCTGAAGATTCTTTGACCAAGTCTCTGGGGCCCTGTTGTGAGAGCAGATT 1976
Db 224 AGTGGATACCTGAAGATTCTTTGACCAAGTCTCTGGGGCCCTGTTGTGAGAGTATT 283
QY 1977 GCTACTAGGCAAGATCTGAGATTGAAAGAGAGAGAGAGATGAGGTGACGCCCA 2036
Db 284 GCTACTAGGCAAGATCTGAGATTGAAAGAGAGAGAGAGATGAGGTGACGCCCA 343
QY 2037 CAACGACCACTCATGGGCTTTGAGAGAGATCTGAAAGAGAGATCTCTCAGGGGGCCAGACC 2096
Db 344 CAGTGACCACTCATGGGCTTTGAGAGAGATCTGAAAGAGAGATCTCTCTGGGGCCAGGCC 403
QY 2097 TGGACTTCCCTATGGGCTGAGTGACACAGATCTGGGGGGCGCGCGCACTAAGTGCGGA 2156
Db 404 TGGGCTTCCCTATGGGCTGAGCGACATGATCTGGGGGGCGCGCGCACTAAGTGCGGA 463
QY 2157 GAGTGAAGTTGAGAACCAACCCAGGGGTCCAGAGGGGCGCCAGGGGTGAGAGGGCCAGGCC 2216
Db 464 GAGTGAAGTTGAGAGAGCCAGCCAGGGGTCCAGAGGGGCGCCAGGGGTGAGAGGGCCAGGCC 523
QY 2217 AGCCTGTGAGTGTGGGGGGCGCGACAGGTGAGGGGCGCGTGTGTGGGGCGAGAGGGG 2276
Db 524 AGCCTGTGAGTGTGGGGGGCGCGACAGGTGAGGGGCGCGTGTGTGGGGCGAGAGGGCC 583
QY 2277 GGGTGGGGGGCGCGCGTGTGGGGGGCGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2336
Db 584 GGGTGGGGGGCGCGCGT 643
QY 2337 GTCCCACTACTCGAGCCACCTGAAAGCGGCACATGCAGACACACAGCGGGGAGAGCCGTT 2396
Db 644 GTCCCACTACTCGAGCCACCTGAAAGCGGCACATGCAGACACACAGCGGGGAGAGCCGTT 703
QY 2397 CCGTGTGGGGCGCGTGTGGGGCGCGTGTGGGGCGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2456
Db 704 CCGTGTGGGGCGCGTGTGGGGCGCGTGTGGGGCGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 763
QY 2457 CACCCATACTGGCGAGAGAGCCCTACCGTGTGTGGGGCGCGTGTGTGTGTGTGTGTGTGTGTGT 2516
Db 764 CACCCATACTGGCGAGAGAGCCCTACCGTGTGTGGGGCGCGTGTGTGTGTGTGTGTGTGTGTGT 823
QY 2517 GGGCAACCTGAGCGGCATCAGCGCACCCACAGAGGCGCTCCACACTCTCTCTCTCTCTCTCTCTCT 2576
Db 824 GGGCAACCTGAGCGGCATCAGCGCACCCACAGAGGCGCGCTCCACACTCTCTCTCTCTCTCTCTCTCT 883
QY 2577 CTGTGGCTTTGATGCTGTCTCCAGACCAACCCGGCTCCACAGTCCCAAGAGAGAGAGAGAGAGAG 2636
Db 884 CTGTGGCTTTGATGCTGTCTCCAGACCAACCCGGCTCCACAGTCCCAAGAGAGAGAGAGAGAGAG 943
QY 2637 GGGGACAAATGCCCGACGATCAGAAATGCGTGTATCTCTGCCAGACTTGAGTCTTCAATGT 2696
Db 944 GGGGGCGGTGCGCGGGCGACCTGAAGATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1003
QY 2697 GCCACCAAGTGTCCAGATTCTCTGCCAGACTGTGGGAGTGTGGGAGTGTGGGAGTGTGGGAGTGTGG 2756
Db 1004 GCCACCAAGTGTCCAGATTCTCTGCCAGACTGTGGGAGTGTGGGAGTGTGGGAGTGTGGGAGTGTGG 1063
QY 2757 CTGTGTGGAATCGATCCGAAACCACTGCCAGAGCTACTGTTCCTCTTGGACCTCTCCCGGG 2816
Db 1064 CTCTGCGGAGTGTGATCAGAACCACTGCCAGAGCTGTCTATTCCTTGGACCTCTCCCGGG 1123
QY 2817 CTGTGCAAGAACTGAGAGAGGTGAGGGCAGAGGCTGGGAGCTGCCATGTGTGGCG 2876
Db 1124 CTGTGCAAGAACTGAGAGAGGTGAGGGTGTGCGGTGTGGGAGCTGCCATGTGTGGCG 1183
QY 2877 CTGCATGCGAGAGAGGCTGAGGGGTGTGCGGTGTGGGAGCTGCCATGTGTGGCG 2936
Db 1184 CTGCATGCGAGAGAGGCTGAGGGGTGTGCGGTGTGGGAGCTGCCATGTGTGGCG 1243
QY 2937 AGGCTTTGCTGTAGTTTATGCCCTTTTGCACCTCACTATACCCCAACCACTGGCTCGGCA 2996
Db 1244 AGGCTTTGCTGTAGTTTATGCCCTTTTGCACCTCACTATACCCCAACCACTGGCTCGGCA 1303

559 CAGCCAGCTCGTCAACTGACACGATACCCGACCCACACTGGCGAGAGCCCTTACC 618
2483 GTTGTCCCACTGCCCCCTTTGCTGTCAGACGCTGGGCAACCTGAGGGGCACTAGCGCA 2542
619 GCTGTCCCACTGCCCCCTTTGCTGTCAGACGCTGGGCAACCTGAGGGGCACTAGCGTA 678
2543 CCCACAGAGGCGCTCCCACTGCTCCCTGCCCAACCTGTGGCTTTGCTGCTGCTCCAC 2602
679 CCACGAGGGGCCCCCACTCTCCCTGCCGACCTGTGGCTTCGGTGTGTACTCCAC 738
2603 GACCAACCCGGGCTCCAGTCCACAGAGCAGGAGGGGCAATGCCGCCGAGATCAGAAA 2662
739 GACCAAGCCGGGCTCCCACTGCTCCCAAGAGCAGGAGGGGCGGTGCCCGCGACCTGAAG 798
2663 ATGGCTCATCTGTCAGACTTGAATCTCATGTGTCACACAGGTGCTGCCAGTTTCTGTC 2722
799 ATGCTCTGCTCTCTCCAGATTTGAGCTTCCATGTGTCACAGAGTGGTGGCAGTTTCTGTC 858
2723 CAGACTGTGGGAGCTGCGGGGTGAAGGGGAGAGTGTGTGGAACCTGGATCCGAAACAC 2782
859 CAGACTGTGGGAGCTGCGGGGTGAAGGGGAGAGGCTCTGCGGGAGCTGGATCAGAAACAC 918
2783 TGCCAGAGCTACTGTTCCCTTTGGACCTGCCGGGCTGTGGACAGCACTGGAGGAGGGTG 2842
919 TGCCAGAGCTGCTATTTCCCTTTGGACCTGCCGGGCTGTGGACAGCACTGGAGGAGGGTG 978
2843 AGGGCAGCAGGCTGGGAGCTGCATGTGTGGGCGCTGTCATCGAGGAGAGCTGGAGGGG 2902
979 AGGGTAGTGGCTGGGAGCTGCATGTGTGGGCGCTGTCATCGAGGAGAGCTGGAGGGG 1038
2903 TTGCCACTGGGGGACCCAGGCGCTGTGTGACAAAGGCTTTGCCGTAGTTTATGCCCT 2962
1039 GTGCAGTGGGGGGCCCCAGGCGCCCGAGTGAACAAAGGCTTTGCCGTAGCTTGGCCCT 1098
2963 TTGCCACTCAGTACCCCAACCACTGGCTGGGCAATGAAGACTCAGAGTGTGTGAGAAC 3022
1099 TTGCCACTCAGTATCCCAACCACTGGCGCGGACATGAGACACACAGTGTGTGAGAAC 1158
3023 CTTTCGGCTGTGCGGCTGTGCATAGGCTGTGTGCTCATCTGATTAACCTGAACCGGAC 3082
1159 CTTTCGGCTGTGCGGCTGTGCTTATGCTGTGTGCTCATCTGATTAACCTGAACCGGAC 1218
3083 ACGCGTCCACAGAGAGAAAGCCCTACAAGTGGCCCTCTGTCCGCTATGCTGTGGCA 3142
1219 ACGGCTCATACAGGAGAGAGCCCTACAAGTGGCCCTCTGTCCCTTATGCTGTGGCA 1278
3143 ACCTGGCCAACTCAAGCTATGCTGTCATCTGCTGTCATCTGCTGTCATCTGCTGCTGTA 3202
1279 ATCTGGCCAACTCAAGCTATGCTGTCATCTGCTGTCATCTGCTGTCATCTGCTGCTGTA 1338
3203 GCTTTGCACTACAGCTGCACAGAGTATGAACCTCAAGCTCATATGCTGTCGACACA 3262
1339 GCTTTGCACTACAGCTGCACAGAGTATGAACCTCAAGCTCATATGCTGTCGACACA 1398
3263 CGGGCAGAGAGCCCTTCCGCTGTGCCACCTGGGCTTATACCAAGGCACTGGGCAACT 3322
1399 CAGGCGAGAGAGCCCTTCCGCTGTGCCACCTGGGCTTATACCAAGGCACTGGGCAACT 1458
3323 ACAGGCTCATCAGAGTGCATGGCCATGTTGGAGCAGAGGCGCTGCTCTCTGCGCC 3382
1459 ACAGGCGCCACCAAGAGTGCATGGCCATGTTGGAGCAGAGGCGCTGCTCTCTGCGCC 1518
3383 CTGAGGCTGGGCGCCCACTCATAGCCCACTCTGTTTGGAGCACTCGGGGTCAGCAG 3442
1519 CTGAGGCTGGGCGCCCACTCATAGCCCACTCTGTTTGGAGCACTCGGGGTCAGCAG 1578
3443 CCTGGGTGCTATGTTAGAGGCTCTTCATTTAGACTCAGCTTGAACCTAGGTTC 3502
1579 CCTGGGAGCTGTGGCAGCGGGCTGTCCACAGACTCATCTGTAAC----- 1627
3503 TTTTACTGGGGCTCTAGGAATTTAGCCCTATGCTCTGCTATTTATACAAATGAACCTAG 3562

1628 -----TAGGTCTTCTTCCCATGTTTATACAGACGACCGAG 1665
3563 AAACACACCTTTCCTTTCCTCCCTGCTCAGGGGCTCCACACAGACTTAACCTAGGCAC 3622
1666 AAGCCACCTT-----TTTCTCCCTGCTGCGCAGGGGCTCCACACAGACTTAACCTAGGCAC 1721
3623 TATATGGACACGCTGAATCCCATGTCAGGGG-CCATATAGACAGGGGA-CTTGCT 3680
1722 TATAAGGACACGCCCAACCCCATGGCGGGGGGCCATATGACACAGGGGACCTTGCT 1781
3681 TAGCTCAGCTTACAGATGAGCTAAGTGTATGAGGCTTGGATTCACCGCACTGCTCCCA 3740
1782 TGACTGAGCACTTACAGAGCTCAGTGAAGAGGCGCTGTATTCACCTCCACTGCCCCA 1841
3741 GAGGCTATGATGAATGCTTGGAG-CTGCCAGCCTTTTAC-----TGTTTTAACTTA 3794
1842 GGGCTGTGGACAAACCGCTGGGGAGCTGCCAGCTCCCTGTTTATTTAACTTAT 1901
3795 TTTCACTGCTTATATATGAAGGAACACTTAACAGAAAAAATAAAAAA 3847
1902 TTCAAGTGTCTTATATATGAAGGAACACTTAACAGAAAAAATAAAAAA 1954

RESULT 9
AAI59942
ID AAI59942 standard; cdna; 1488 BP.
XX AAI59942;
AC AAI59942;
XX
DT 22-OCT-2001 (first entry)
XX
XX Human polynucleotide SEQ ID NO 3931.
DE
XX Human; nontropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia; ss.
XX
XX Homo sapiens.
OS
XX WO200153312-A1.
XX
XX 26-JUL-2001.
PD
XX 26-DEC-2000; 2000WO-US034263.
PF
XX 23-DEC-1999; 99US-00471275.
PR 21-JAN-2000; 2000US-00488725.
PR 25-APR-2000; 2000US-00552317.
PR 20-JUN-2000; 2000US-00598042.
PR 19-JUL-2000; 2000US-00620312.
PR 03-AUG-2000; 2000US-00653450.
PR 14-SEP-2000; 2000US-00662191.
PR 19-OCT-2000; 2000US-00693036.
PR 29-NOV-2000; 2000US-00727344.
XX
XX (HYSE-) HYSEQ INC.
PA
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
PI Zhou P, Goodrich R, Drmanac R;
XX WPI; 2001-442253/47.
DR P-PSDB; AAM40786.
XX
XX Novel nucleic acids and polypeptides, useful for treating disorders such
PT as central nervous system injuries.
XX
XX Claim 1; SEQ ID NO 3931; 10078pp; English.
XX
XX The invention relates to human nucleic acids (AAI57798-AAI61369) and the

CC encoded polypeptides (AAM38642-AM42213) with nootropic,
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemias and
 CC C.N.S disorders. Note: The sequence data for this patent did not form
 CC part of the printed specification

SQ Sequence 1488 BP; 287 A; 460 C; 484 G; 257 T; 0 U; 0 Other;

Query Match	31.7%;	Score 1220.2;	DB 4;	Length 1488;
Best Local Similarity	92.9%;	Pred. No. 4.1e-301;		
Matches 1272;	Conservative 0;	Mismatches 97;	Indels 0;	Gaps 0;
QY 1885	GAGGAGATTCCAAAGCACCCTCCCTCTTTAGTGATCTAGAGATTCTCTCGACGAA	1944		
DB 119	GAGGAGATTCCAAAGCACCCTCCCTCTTTAGTGATCTAGAGATTCTCTCGACGAA	178		
QY 1945	GGTCTGGGGCCCTGGTGTGGAGAGCGATTGCTATAGGCCAAGATCTGGAGTTGAA	2004		
DB 179	GGACCCGGGGCCCTGGTATTGGAGAGTGAATTGCTATAGGCCAAGATCTGGAGTTGAG	238		
QY 2005	GAGGAGAGGAGGAGGATGAGGTGAGCGGCACACACAGCAGCTCATGGGCTTTGAGAG	2064		
DB 239	GAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	238		
QY 2065	GACTCTGAGGAGACTCTCAGGGGCCAGACCTGGACTTCCCTATGGGCTGAGTGACGAC	2124		
DB 299	GACTCGGAGGAGACTCTCTGGGGGCCAGGCTGGGCTTCCCTATGGGCTGAGCGACGAT	358		
QY 2125	GAGTCTGGGGGCGCGCGCACTAAGTCCGAGAGTGAAGTTGAGGACCGACGAGGGGT	2184		
DB 359	GAGTCTGGGGGCGCGCGCACTAAGTCCGAGAGTGAAGTTGAGGAGCGACGAGGGGT	418		
QY 2185	CCAGGGAGGGCCAGGGGTGAGAGGCCAGGCCAGCCTGTCTGAGTGTGTGGGGGGCCGACA	2244		
DB 419	CCAGGGAGGGCCAGGGGTGAGAGGCCAGGCCAGCCTGTCTGAGTGTGTGGGGGGCCGACA	478		
QY 2245	GGTGGGGGGCGGTGTGGGGGCGAGGGGCGGGTGGGGGGGGGGGGGGGGGGGGGGGGGG	2304		
DB 479	GGTGGGGGGCGGTGTGGGGGCGAGGGGCGGGTGGGGGGGGGGGGGGGGGGGGGGGGGG	538		
QY 2305	TTACTGTACTCATGCGGGCTGTGGGCTTTCGTGTCCTGCTCCACTACTCGAGCCACTGAGCGG	2364		
DB 539	CTACTGTACTCATGCGGGCTGTGCACTTCTGTCCTGCTCCACTACTCGAGCCACTGAGCGG	598		
QY 2365	CACATGCGAGACACAGCGGGGAGAGCGGTTCGGCTGTGGCGCGCTGCCCATACGCTCA	2424		
DB 599	CACATGCGAGACACAGCGGGGAGAGCGGTTCGGCTGTGGCGCGCTGCCCATACGCTCA	658		
QY 2425	GCCAGGTTGTCACCTGACGCGCATATACCGCACCCACTACTGCGAGAGCCCTACCGT	2484		
DB 659	GCCAGGTTGTCACCTGACGCGCATATACCGCACCCACTACTGCGAGAGCCCTACCGC	718		
QY 2485	TGTCCCACTGCCCCCTTTGCTGTGAGAGAGCTGGGCAACCTGAGCGGCACTACGCGACC	2544		
DB 719	TGTCCCACTGCCCCCTTTGCTGTGAGAGAGCTGGGCAACCTGAGCGGCACTACGCGTACC	778		
QY 2545	CACAGAGGGCTCCCACTCCCTCCCTGCGGCAACCTGCTGCTTTCGATGCTGCTCCACGA	2604		
DB 779	CACAGAGGGCTCCCACTCCCTCCCTGCGGCAACCTGCTGCTTTCGATGCTGCTCCACGA	838		
QY 2605	CCAAACCGGCTCCAGTCCCAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	2664		
DB 839	CCAAACCGGCTCCAGTCCCAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	898		

QY 2665	GCGTGATCTCTCCAGAGCTTCTGAGTCTTCTATGTGCACACAGGTGGTCCAGTTCTCTGCCA	2724		
DB 899	GCTCTGCTCTCTTCCAGATTTCAGGCTTCCATGTGCCACAGGTGGTGCAGTTTCTTGCCA	958		
QY 2725	GACTGTGGGAGCTGGCGGGTGAAGGGAGAGCTTGTGTGGAATCTGATCCGAACCACTG	2784		
DB 959	GACTGTGGGAGCTGGCGGGTGAAGGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	1018		
QY 2785	CCAGAGCTTACTTGTTCCTTGGACCTGCGCGGGCTGTGACAGGAACTGAGAGGAGGTGAG	2844		
DB 1019	CCAGAGCTTACTTATTCCTTGGACCTGCGCGGGCTGTGACAGGAGCTGAGAGGAGGTGAG	1078		
QY 2845	GGCAGAGGCTGGGAGCTGCGATGTGTGGGCGCTGATGCGAGAGAGCTGGAGGGGTT	2904		
DB 1079	GCTAGTCTGGGAGCTGCGATGTGTGGGCGCTGATGCGAGAGAGCTGGAGGGGTT	1138		
QY 2905	GCCACTGGGGGACCCCGAGGCGCTGTGACAAAGGCTTGTGCTGTAGTTATGCCCTTT	2964		
DB 1139	GCCAGTGGGGGCGCCCGAGGCGCTGTGACAAAGGCTTGTGCTGTAGTTATGCCCTTT	1198		
QY 2965	GCCACTCTACTACCCCAACACCTGGCTCGGCACATGAAAGATCTCAGTGTGAGAGAACCC	3024		
DB 1199	GCCACTCTACTATCCCAACACCTGGCTCGGCACATGAAAGATCTCAGTGTGAGAGAACCC	1258		
QY 3025	TTCCGCTGTGCGGCTGTCCATACGCTCTGCTCATCTGATGATACCTGAAACGGCACCG	3084		
DB 1259	TTCCGCTGTGCGGCTGTCCATACGCTCTGCTCATCTGATGATACCTGAAACGGCACCG	1318		
QY 3085	CGCGTCCACACAGGAGAGAAAGCCCTTACAGTGGCCCTCTGTCGATGCTGCTGGCAAC	3144		
DB 1319	CGCGTCCACACAGGAGAGAGAGCCCTTACAGTGGCCCTCTGTCGATGCTGCTGGCAAT	1378		
QY 3145	CTGGCCCACTCAAGGCTGATGGTGGATCCACTCTGCTGATCAAACTTTCGGTGTAGC	3204		
DB 1379	CTGGCCCACTCAAGGCTGATGGTGGATCCACTCTGCTGATCAAACTTTCGGTGTAGC	3264		
QY 3205	CTTTGCACTACAGCTGCAACCCAGAGATGAACTTCAAACTGCAATATGC	3253		
DB 1439	CTTTGCACTACAGCTGCAACCCAGAGATGAACTTCAAACTGCAATATGC	1487		
RESULT 10	AAK85394			
ID	AAK85394	standard; DNA; 9266 BP.		
XX	AAK85394;			
AC	AAK85394;			
XX	07-NOV-2001 (first entry)			
DT	07-NOV-2001			
XX	Human immune/haematopoietic antigen genomic sequence SEQ ID NO:40206.			
DE	Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;			
XX	cytostatic; gene therapy; vaccine; metastasis; ds.			
KW	Homo sapiens.			
OS	WO200157182-A2.			
XX	09-AUG-2001.			
XX	17-JAN-2001; 2001WO-US001354.			
PF	31-JAN-2000; 2000US-0179065P.			
XX	04-FEB-2000; 2000US-0180628P.			
PR	24-FEB-2000; 2000US-0184664P.			
PR	02-MAR-2000; 2000US-0186350P.			
PR	16-MAR-2000; 2000US-0189874P.			
PR	17-MAR-2000; 2000US-0190076P.			
PR	18-APR-2000; 2000US-0198123P.			
PR	19-MAY-2000; 2000US-0205515P.			
PR	07-JUN-2000; 2000US-0209467P.			
PR	28-JUN-2000; 2000US-0214886P.			
PR	30-JUN-2000; 2000US-0215135P.			

PR 07-JUL-2000; 2000US-02116647P.
PR 07-JUL-2000; 2000US-02116880P.
PR 11-JUL-2000; 2000US-02117480P.
PR 11-JUL-2000; 2000US-02117487P.
PR 11-JUL-2000; 2000US-02117496P.
PR 14-JUL-2000; 2000US-02118290P.
PR 26-JUL-2000; 2000US-0220863P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-02252113P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241121P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.

PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0241617P.
PR 01-NOV-2000; 2000US-0241647P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 17-NOV-2000; 2000US-024613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-483426/52.
XX
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
XX useful for preventing, diagnosing and/or treating cancers and metastasis.
XX
XX Disclosure; SEQ ID NO 40206; 3071pp + Sequence Listing; English.
XX
XX AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
XX amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
XX activity, and can be used in gene therapy and vaccine production. (I)
XX proteins and polynucleotides may be used in the prevention, diagnosis and
XX treatment of diseases associated with inappropriate (I) expression. For
XX example, they may be used to treat disorders associated with decreased
XX expression by rectifying mutations or deletions in a patient's genome
XX that affect the activity of (I) by expressing inactive proteins or to
XX supplement the patients own production of (I). Additionally, (I)

RESULT 11
AAL04937
ID AAL04937 standard; DNA; 9266 BP.
XX AC AAL04937;
XX DT 21-NOV-2001 (first entry)
XX DE Human reproductive system related antigen DNA SEQ ID NO: 7625.
XX KW Human; reproductive system related antigen; reproductive system disorder;
XX KW cancer; gene therapy; ds.
XX OS Homo sapiens.
XX PN WC200155320-A2.
XX PD 02-AUG-2001.
XX PF 17-JAN-2001; 2001WO-US001339.
XX PR 31-JAN-2000; 2000US-0179065P.
XX PR 04-FEB-2000; 2000US-0180628P.
XX PR 24-FEB-2000; 2000US-0184664P.
XX PR 02-MAR-2000; 2000US-0186350P.
XX PR 16-MAR-2000; 2000US-0189874P.
XX PR 17-MAR-2000; 2000US-0190076P.
XX PR 18-APR-2000; 2000US-0198123P.
XX PR 19-MAY-2000; 2000US-0205515P.
XX PR 07-JUN-2000; 2000US-0209467P.
XX PR 28-JUN-2000; 2000US-021486P.
XX PR 30-JUN-2000; 2000US-0215135P.
XX PR 07-JUL-2000; 2000US-0216647P.
XX PR 07-JUL-2000; 2000US-0216880P.
XX PR 11-JUL-2000; 2000US-0217487P.
XX PR 11-JUL-2000; 2000US-0217496P.
XX PR 14-JUL-2000; 2000US-0220963P.
XX PR 26-JUL-2000; 2000US-0220964P.
XX PR 26-JUL-2000; 2000US-0220964P.
XX PR 14-AUG-2000; 2000US-0224518P.
XX PR 14-AUG-2000; 2000US-0224519P.
XX PR 14-AUG-2000; 2000US-0225213P.
XX PR 14-AUG-2000; 2000US-0225214P.
XX PR 14-AUG-2000; 2000US-0225266P.
XX PR 14-AUG-2000; 2000US-0225267P.
XX PR 14-AUG-2000; 2000US-0225268P.
XX PR 14-AUG-2000; 2000US-0225270P.
XX PR 14-AUG-2000; 2000US-0225447P.
XX PR 14-AUG-2000; 2000US-0225757P.
XX PR 14-AUG-2000; 2000US-0225758P.
XX PR 14-AUG-2000; 2000US-0225759P.
XX PR 18-AUG-2000; 2000US-0226273P.
XX PR 22-AUG-2000; 2000US-0226681P.
XX PR 22-AUG-2000; 2000US-0226686P.
XX PR 22-AUG-2000; 2000US-0227182P.
XX PR 23-AUG-2000; 2000US-0227009P.
XX PR 30-AUG-2000; 2000US-0228924P.
XX PR 01-SEP-2000; 2000US-0229287P.
XX PR 01-SEP-2000; 2000US-0229343P.
XX PR 01-SEP-2000; 2000US-0229344P.
XX PR 01-SEP-2000; 2000US-0229345P.
XX PR 05-SEP-2000; 2000US-0229509P.
XX PR 06-SEP-2000; 2000US-0229513P.
XX PR 06-SEP-2000; 2000US-0230437P.
XX PR 06-SEP-2000; 2000US-0230438P.
XX PR 08-SEP-2000; 2000US-0231243P.
XX PR 08-SEP-2000; 2000US-0231243P.
XX PR 08-SEP-2000; 2000US-0231244P.
XX PR 08-SEP-2000; 2000US-0231413P.
XX PR 08-SEP-2000; 2000US-0231414P.
XX PR 08-SEP-2000; 2000US-0233080P.
XX PR 08-SEP-2000; 2000US-0233080P.
XX PR 08-SEP-2000; 2000US-0233081P.
XX PR 12-SEP-2000; 2000US-0233198P.
XX PR 14-SEP-2000; 2000US-0233297P.
XX PR 14-SEP-2000; 2000US-0233298P.
XX PR 14-SEP-2000; 2000US-0233399P.
XX PR 14-SEP-2000; 2000US-023400P.
XX PR 14-SEP-2000; 2000US-023401P.
XX PR 14-SEP-2000; 2000US-0233063P.
XX PR 14-SEP-2000; 2000US-0233064P.
XX PR 14-SEP-2000; 2000US-0233065P.
XX PR 21-SEP-2000; 2000US-0234223P.
XX PR 21-SEP-2000; 2000US-0234274P.
XX PR 25-SEP-2000; 2000US-0234997P.
XX PR 25-SEP-2000; 2000US-0234998P.
XX PR 26-SEP-2000; 2000US-0235484P.
XX PR 27-SEP-2000; 2000US-0235834P.
XX PR 27-SEP-2000; 2000US-0235836P.
XX PR 29-SEP-2000; 2000US-0236327P.
XX PR 29-SEP-2000; 2000US-0236367P.
XX PR 29-SEP-2000; 2000US-0236368P.
XX PR 29-SEP-2000; 2000US-0236369P.
XX PR 29-SEP-2000; 2000US-0236370P.
XX PR 02-OCT-2000; 2000US-0236802P.
XX PR 02-OCT-2000; 2000US-0237037P.
XX PR 02-OCT-2000; 2000US-0237038P.
XX PR 02-OCT-2000; 2000US-0237039P.
XX PR 13-OCT-2000; 2000US-0237040P.
XX PR 13-OCT-2000; 2000US-0239335P.
XX PR 13-OCT-2000; 2000US-0239337P.
XX PR 20-OCT-2000; 2000US-0240960P.
XX PR 20-OCT-2000; 2000US-0241221P.
XX PR 20-OCT-2000; 2000US-0241785P.
XX PR 20-OCT-2000; 2000US-0241786P.
XX PR 20-OCT-2000; 2000US-0241787P.
XX PR 20-OCT-2000; 2000US-0241808P.
XX PR 20-OCT-2000; 2000US-0241809P.
XX PR 20-OCT-2000; 2000US-0244617P.
XX PR 01-NOV-2000; 2000US-0244617P.
XX PR 08-NOV-2000; 2000US-0246474P.
XX PR 08-NOV-2000; 2000US-0246475P.
XX PR 08-NOV-2000; 2000US-0246476P.
XX PR 08-NOV-2000; 2000US-0246477P.
XX PR 08-NOV-2000; 2000US-0246478P.
XX PR 08-NOV-2000; 2000US-0246523P.
XX PR 08-NOV-2000; 2000US-0246524P.
XX PR 08-NOV-2000; 2000US-0246525P.
XX PR 08-NOV-2000; 2000US-0246526P.
XX PR 08-NOV-2000; 2000US-0246527P.
XX PR 08-NOV-2000; 2000US-0246528P.
XX PR 08-NOV-2000; 2000US-0246532P.
XX PR 08-NOV-2000; 2000US-0246609P.
XX PR 08-NOV-2000; 2000US-0246610P.
XX PR 08-NOV-2000; 2000US-0246611P.
XX PR 08-NOV-2000; 2000US-0246613P.
XX PR 17-NOV-2000; 2000US-0249207P.
XX PR 17-NOV-2000; 2000US-0249208P.
XX PR 17-NOV-2000; 2000US-0249209P.
XX PR 17-NOV-2000; 2000US-0249210P.
XX PR 17-NOV-2000; 2000US-0249211P.
XX PR 17-NOV-2000; 2000US-0249212P.
XX PR 17-NOV-2000; 2000US-0249213P.
XX PR 17-NOV-2000; 2000US-0249214P.
XX PR 17-NOV-2000; 2000US-0249215P.
XX PR 17-NOV-2000; 2000US-0249216P.
XX PR 17-NOV-2000; 2000US-0249217P.
XX PR 17-NOV-2000; 2000US-0249218P.
XX PR 17-NOV-2000; 2000US-0249244P.
XX PR 17-NOV-2000; 2000US-0249245P.
XX PR 17-NOV-2000; 2000US-0249264P.
XX PR 17-NOV-2000; 2000US-0249265P.
XX PR 17-NOV-2000; 2000US-0249297P.
XX PR 17-NOV-2000; 2000US-0249299P.
XX PR 17-NOV-2000; 2000US-0249300P.

3022	TCCTTCTTCCCATGTTTATATACAGCGGACCAAGCCACCTT-----TTTCTCCCCGC	3077
3589	TGGTCAGGGGCTCCACACAGACTTAACCTAGGCACTATATGGACCAGCCTGAATCCCATGG	3648
3078	TGGCAGGGGCTCCACACAGACTTAACCTAGGCACTATATAGGACCAGCCCAACCCCATGGG	3137
3649	TCAGGGGG--CCATATAGACCCAGGGGA--CTTGTCTTTAGCTCAGTACCAGATGAGCTAAGT	3706
3138	CGGGGGGGCCCATATGAGACCCAGGGGACCTTGCCTTGACTGAGGCACTTCACGAGCTCAGT	3197
3707	GATTAGGGCTTTGGATTCAACCGCACTGCTCCAGAGGCTATGGATGAACCTGTTGGGAG	3766
3198	GAGAAGGGCCCTGTATTACCTCCACTGCCCCAGGGGCTGTGGCAAAACCGGCTGGGGG	3257
3767	-CTGCCCCAGCCTTTTAC-----TGTTTTAACTTATTTTTCAGTTCGCTTTATATAAAGGAAACA	3821
3258	ACTGCCACGCTCCCACTCTTTATTAACTTATTTTTCAGTTCGCTTTATATAAAGGAAACA	3317
3822	CTAACAGA	3829
3318	CTAACAAA	3325
RESULT 12		
ABL97831	standard; DNA; 9266 BP.	
ID	ABL97831	
XX	AC	
XX	ABL97831;	
21-JUN-2002	(first entry)	
DE	Human testicular antigen encoding DNA fragment SEQ ID NO: 2483.	
XX	Human; testicular antigen; testes; cancer; metastasis; immune disorder;	
KW	reproductive system disorder; urinary system disorder; gene therapy;	
KW	cardiovascular disorder; respiratory disorder; neurological disorder;	
KW	gastrointestinal disease; infection; cytostatic; gene; ds.	
XX	Homo sapiens.	
OS	Homo sapiens.	
XX	WO200155317-A2.	
PN	WO200155317-A2.	
PD	02-AUG-2001.	
XX	17-JAN-2001; 2001WO-US0001329.	
PF	31-JAN-2000; 2000US-0179065P.	
XX	04-FEB-2000; 2000US-0180628P.	
XX	24-FEB-2000; 2000US-0184664P.	
XX	02-MAR-2000; 2000US-0186350P.	
XX	16-MAR-2000; 2000US-0189874P.	
XX	17-MAR-2000; 2000US-0190076P.	
XX	18-APR-2000; 2000US-0198123P.	
XX	19-MAY-2000; 2000US-0205515P.	
XX	07-JUN-2000; 2000US-0209467P.	
XX	28-JUN-2000; 2000US-0214886P.	
XX	30-JUN-2000; 2000US-0215135P.	
XX	07-JUL-2000; 2000US-0216647P.	
XX	07-JUL-2000; 2000US-0216800P.	
XX	11-JUL-2000; 2000US-0217487P.	
XX	11-JUL-2000; 2000US-0217496P.	
XX	14-JUL-2000; 2000US-0218290P.	
XX	26-JUL-2000; 2000US-0220963P.	
XX	26-JUL-2000; 2000US-0220964P.	
XX	14-AUG-2000; 2000US-0224518P.	
XX	14-AUG-2000; 2000US-0224519P.	
XX	14-AUG-2000; 2000US-0225213P.	
XX	14-AUG-2000; 2000US-0225214P.	
XX	14-AUG-2000; 2000US-0225266P.	
XX	14-AUG-2000; 2000US-0225267P.	
XX	14-AUG-2000; 2000US-0225268P.	
XX	14-AUG-2000; 2000US-0225270P.	

PR 08-NOV-2000; 2000US-0246532P.
 PR 08-NOV-2000; 2000US-0246603P.
 PR 08-NOV-2000; 2000US-0246610P.
 PR 08-NOV-2000; 2000US-0246611P.
 PR 08-NOV-2000; 2000US-0246613P.
 PR 17-NOV-2000; 2000US-0249207P.
 PR 17-NOV-2000; 2000US-0249208P.
 PR 17-NOV-2000; 2000US-0249209P.
 PR 17-NOV-2000; 2000US-0249210P.
 PR 17-NOV-2000; 2000US-0249211P.
 PR 17-NOV-2000; 2000US-0249212P.
 PR 17-NOV-2000; 2000US-0249213P.
 PR 17-NOV-2000; 2000US-0249214P.
 PR 17-NOV-2000; 2000US-0249215P.
 PR 17-NOV-2000; 2000US-0249216P.
 PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249244P.
 PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249264P.
 PR 17-NOV-2000; 2000US-0249265P.
 PR 17-NOV-2000; 2000US-0249297P.
 PR 17-NOV-2000; 2000US-0249298P.
 PR 17-NOV-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 05-DEC-2000; 2000US-0256719P.
 PR 06-DEC-2000; 2000US-0251473P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 08-DEC-2000; 2000US-0251990P.
 PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-0255678P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Barash SC, Ruben SM;
 XX
 XX WPI; 2001-483232/52.
 XX
 DR Nucleic acids encoding 973 human testicular antigen polypeptides, useful
 PT for preventing, diagnosing and/or treating testicular cancer.
 XX
 PS Disclosure; SEQ ID NO 2483; 766pp; English.
 XX
 CC The present invention provides the protein and coding sequences of 973
 CC human testicular antigens, and fragments of their genomic sequences. The
 CC sequences can be used in the treatment of cardiovascular, urinary system,
 CC reproductive system, immune, respiratory, neurological and
 CC gastrointestinal disorders, infections, and particularly cancer,
 CC especially testicular cancers. The present sequence is a DNA encoding a
 CC protein fragment of the invention
 XX
 SQ Sequence 9266 BP; 2324 A; 2431 C; 2502 G; 2009 T; 0 U; 0 Other;
 Query Match 31.1%; Score 1195; DB 4; Length 9266;
 Best Local Similarity 81.3%; Pred. No. 2.9e-294; Indels 144; Gaps 10;
 Matches 1568; Conservative 0; Mismatches 216;
 QY 2003 AAGAGGAGAGAGAGAGATGAAGTGACGGGCCACACAGCAGCTCATGGCTTTGAGA 2062
 DB 1441 AAGTGAAAGGGAGGTTGGAGCGCTGATCTTACCTATGCTGATGCTCTTTATGT 1500
 QY 2063 GAGACTCTGAGAGAGACTCTCAGGGGGCCAGACTTGACTTCCCTATGGCTGAGTGACG 2122
 DB 1501 CTATTTACAGAGAGACTCTCTGGGGGCCAGCGCTTCCCTATGGCTGAGCGACG 1560
 QY 2123 ACGAGTCTGGGGGGCCCGCGCACTAAGTGGCGAGAGTGAAGTTGAGGAACACAGCCAGGG 2182

DB 1561 ATGAGTCTGGGGGGCGCGGCACTAAGTGGCGAGAGTGAAGTTGAGGAGCGAGCCAGGG 1620
 QY 2183 GTCCAGGGGAGGCCAGGGGTGAGAGGCCAGGCCAGCCCTGTGTGAGTGTGTGGGGGCGGA 2242
 DB 1621 GTCCAGGGGAGGCCAGGGGTGAGAGGCCAGGCCAGCCCTGTGTGAGTGTGTGGGGGCGGA 1680
 QY 2243 CAGGTGAGGGGCGGTGTTGTGGGGCAGGAGGCGGGGTGGGGGGGCCCCGCTGCCCCCGCAC 2302
 DB 1681 CAGGTGAGGGGCGGTGTTGTGGGGCAGGAGGCGGGGTGGGGGGG-CCCTGTGTCCTCCAC 1739
 QY 2303 GGTACTGTACTATGCGCGCTTGTGGCTTTCGTGTCCCACTACTCGAGCCACCTGGAAGC 2362
 DB 1740 GGCTACTGTACTATGCGCGCTTGTGGCTTTCGTGTCCCACTACTCGAGCCACCTGGAAGC 1799
 QY 2363 GGCACATGACAGACACACAGCGGGGAGAGCCGTTCGCTGTGGCGCTGCGCTGCGCTTACGNT 2422
 DB 1800 GGCACATGACAGACACACAGCGGGGAGAGCCGTTCGCTGTGGCGCTGCGCTTACGCT 1859
 QY 2423 CAGCCAGTTCGTAACTGTGACGGGACATACCCGACCCCATACTGGCGGAGAGCCCTTACC 2482
 DB 1860 CAGCCAGTTCGTAACTGTGACGGGACATACCCGACCCCATACTGGCGGAGAGCCCTTACC 1919
 QY 2483 GTTGTCCCGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTT 2542
 DB 1920 GTTGTCCCGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTT 1979
 QY 2543 CCACACAGGGGCTTCCACTTCTTCCCTGCCCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2602
 DB 1980 CCACACAGGGGCTTCCACTTCTTCCCTGCCCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2039
 QY 2603 GACCAACCGGGCTTCCAGTTCCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2651
 DB 2040 GACCAACCGGGCTTCCAGTTCCTCAGTTCCTCAGTTCCTCAGTTCCTCAGTTCCTCAGTTCCT 2099
 QY 2652 ----- 2651
 DB 2100 TAAGACACACAGGAGCAAGATCTTGGGACATGGGTGGTGCCTAGGAATGCTTGG 2159
 QY 2652 -----ACGATCAGAAATGGCTGATCTCTGCCAGCTTGTGCTTCCAGATTGAGC 2688
 DB 2160 ATTGGATTATAGCCAGGTCTTTGTGCCACAGATGCTGTCTCTTCCAGATTGAGC 2219
 QY 2689 CTTTATGTGTCACAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2748
 DB 2220 CTTTATGTGTCACAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2278
 QY 2749 GGGGAGAGCTTGTGTGAACTGGATCCGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2808
 DB 2279 GGGGAGAGCTTGTGTGAACTGGATCCGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2338
 QY 2809 TGCCGGGCTTGTGACAGGAACTGGAGGAGGTGAGGCGAGCAGGCTGGGAGCTGGCATG 2868
 DB 2339 TGCCGGGCTTGTGACAGGAACTGGAGGAGGTGAGGCGAGCTGGGAGCTGGCATG 2398
 QY 2869 TGTGGGCTTGTGACAGGAGGCTGAGGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2928
 DB 2399 TGTGGGCTTGTGACAGGAGGCTGAGGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2458
 QY 2929 GGTGACAAAGGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2988
 DB 2459 -----ACAAAGGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2514
 QY 2989 GCTCGGCACATGAAGACTCAGAGTGTGTGAGAACTTCCGCTGTGTGCTGCTGCTGCTGCTGCTGCT 3048
 DB 2515 GCCCGGCACATGAAGACACACAGTGTGTGAGAACTTCCGCTGTGTGCTGCTGCTGCTGCTGCTGCT 2574
 QY 3049 GCCTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3108
 DB 2575 GCCTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2634
 QY 3109 TACAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3168
 DB 2635 TACAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2694

PD 16-AUG-2001.

XX	17-JAN-2001;	2001WO-US0011334.
PP		
XX		
PP	31-JAN-2000;	2000US-01790055P.
PP	04-FEB-2000;	2000US-01806238P.
PP	24-FEB-2000;	2000US-0184664P.
PP	02-MAR-2000;	2000US-0186350P.
PP	16-MAR-2000;	2000US-0189874P.
PP	17-NAR-2000;	2000US-0190076P.
PP	18-APR-2000;	2000US-01981123P.
PP	18-MAY-2000;	2000US-0205515P.
PP	07-JUN-2000;	2000US-0209467P.
PP	28-JUN-2000;	2000US-0214886P.
PP	30-JUN-2000;	2000US-0215195P.
PP	07-JUL-2000;	2000US-0216647P.
PP	07-JUL-2000;	2000US-0216880P.
PP	11-JUL-2000;	2000US-0217487P.
PP	11-JUL-2000;	2000US-0217496P.
PP	14-JUL-2000;	2000US-0218290P.
PP	26-JUL-2000;	2000US-0220963P.
PP	26-JUL-2000;	2000US-0220964P.
PP	14-AUG-2000;	2000US-0224518P.
PP	14-AUG-2000;	2000US-0224519P.
PP	14-AUG-2000;	2000US-0225213P.
PP	14-AUG-2000;	2000US-0225214P.
PP	14-AUG-2000;	2000US-0225266P.
PP	14-AUG-2000;	2000US-0225267P.
PP	14-AUG-2000;	2000US-0225268P.
PP	14-AUG-2000;	2000US-0225270P.
PP	14-AUG-2000;	2000US-0225474P.
PP	14-AUG-2000;	2000US-0225757P.
PP	14-AUG-2000;	2000US-0225758P.
PP	18-AUG-2000;	2000US-0225759P.
PP	18-AUG-2000;	2000US-0226279P.
PP	22-AUG-2000;	2000US-0226681P.
PP	22-AUG-2000;	2000US-0226868P.
PP	22-AUG-2000;	2000US-0227182P.
PP	23-AUG-2000;	2000US-0227003P.
PP	23-AUG-2000;	2000US-0228924P.
PP	01-SEP-2000;	2000US-0229287P.
PP	01-SEP-2000;	2000US-0229343P.
PP	01-SEP-2000;	2000US-0229344P.
PP	01-SEP-2000;	2000US-0229345P.
PP	03-SEP-2000;	2000US-0229509P.
PP	05-SEP-2000;	2000US-0229513P.
PP	06-SEP-2000;	2000US-0230437P.
PP	06-SEP-2000;	2000US-0230438P.
PP	08-SEP-2000;	2000US-0231242P.
PP	08-SEP-2000;	2000US-0231243P.
PP	08-SEP-2000;	2000US-0231244P.
PP	08-SEP-2000;	2000US-0231413P.
PP	08-SEP-2000;	2000US-0231414P.
PP	08-SEP-2000;	2000US-0232080P.
PP	08-SEP-2000;	2000US-0232081P.
PP	13-SEP-2000;	2000US-0231968P.
PP	14-SEP-2000;	2000US-0232397P.
PP	14-SEP-2000;	2000US-0232398P.
PP	14-SEP-2000;	2000US-0232399P.
PP	14-SEP-2000;	2000US-0232400P.
PP	14-SEP-2000;	2000US-0232401P.
PP	14-SEP-2000;	2000US-0233063P.
PP	14-SEP-2000;	2000US-0233064P.
PP	21-SEP-2000;	2000US-0233065P.
PP	21-SEP-2000;	2000US-0234223P.
PP	21-SEP-2000;	2000US-0234274P.
PP	21-SEP-2000;	2000US-0234597P.
PP	25-SEP-2000;	2000US-0234598P.
PP	26-SEP-2000;	2000US-0235484P.
PP	27-SEP-2000;	2000US-0235834P.
PP	27-SEP-2000;	2000US-0235836P.
PP	29-SEP-2000;	2000US-0236327P.
PP	29-SEP-2000;	2000US-0236367P.
PP	29-SEP-2000;	2000US-0236368P.

PT	Nucleic acids encoding 3224 human nervous system antigen polypeptides, useful for preventing, diagnosing and/or treating nervous system cancers and metastases.
PT	
PT	
XX	Disclosure, SEQ ID NO 8208; 1701pp + Sequence Listing; English.
XX	
CC	The invention relates to novel genes (ABAI1004-ABA21534) and proteins (ABBI14678-ABBI18001) useful for preventing, treating or ameliorating medical conditions e.g. by protein or gene therapy. The genes are isolated from a range of human tissues disclosed in the specification. The nucleic acids, proteins, antibodies and (ant)agonists are useful in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and ovarian cancer and other cancers of the adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune disorders e.g. Addison's disease, allergies, autoimmune haemolytic anaemia, autoimmune rheumatoid arthritis, diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c) cardiovascular disorders such as myocardial ischaemia; (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f) infectious diseases such as viral, bacterial, fungal and parasitic infections. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX	
SQ	Sequence 9266 BP; 2009 A; 2502 C; 2431 G; 2324 T; 0 U; 0 Other;
Query Match 31.1%; Score 1195; DB 5; Length 9266;	
Best Local Similarity 81.3%; Pred. No. 2.9e-294;	
Matches 1568; Conservativity 0; Mismatches 216; Indels 144; Gaps 10;	
QY	2003 AAGAGGAGAGGAGGAGGATGAAGTACGGGCGCACAACGACGAGCTCATGGCTTTGAGA 2062
DB	7826 AAGTGAAGAGGAGAGTGGAGCTGATCTACCTATCTGATGTCTCTTCTATGT 7767
QY	2063 GAGACTCTGAAGAGAGACTCTCAGGGGGCCAGACCTGGAGCTTCCCTATGGGCTGAGTACG 2122
DB	7766 CTATTTCAACAGAGAGACTCTCTGGGGCCAGGCTTCCCTATGGGCTGAGTACG 7707
QY	2123 ACGAGTCTGGGGCGCGCGCACTAAGTGGCGAGAGTGAAGTTGAGGAACAGCAGG 2182
DB	7706 ATGAGTCTGGGGCGCGCGCACTAAGTGGCGAGTGAAGTTGAGGAACAGCAGG 7647
QY	2183 GTCCAGGGAGGCGCAGGGGTGAGAGCCAGGCGCCAGCTGTGAGTGTGTGGGGGCGCA 2242
DB	7646 GTCCAGGGAGGCGCAGGGGTGAGAGCCAGGCGCCAGCTGTGAGTGTGTGGGGGCGCA 7587
QY	2243 CAGGTGAGGGCGCGTGTGTGGGGCGAGGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 2302
DB	7586 CAGGTGAGGGCGCGTGTGTGGGGCGAGGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 7528
QY	2303 GGTACTGTACTCATGCGCGCTGTGGCGTTTGTGTCCTCACTACTCGACCCACCTGAAGC 2362
DB	7527 GGCTACTGTACTCATGCGCGCTGTGACCTCTGTCCTGTCCTACTCGAGCCACCTGAAGC 7468
QY	2363 GGCAATGCGAGACACAGCGGGGAGAGCGCTTCGGCTGTGGCGCTGCCATAGGNT 2422
DB	7467 GGCAATGCGAGACACAGCGGGGAGAGCGCTTCGGCTGTGGCGCTGCCATAGGNT 7408
QY	2423 CAGCCACAGTTGCTCAACCTGACGCGGACATACCCGACCCCATACTGGCGAGAGCCCTACC 2482
DB	7407 CAGCCACAGTTGCTCAACCTGACGCGGACATACCCGACCCCATACTGGCGAGAGCCCTACC 7348
QY	2483 GTTGTCCCACTGCCCTTTGCTGAGAGAGCTGGGCAACCTGAGGGGGCATCAGGCA 2542
DB	7347 GCTGTCCCACTGCCCTTTGCTGAGAGAGCTGGGCAACCTGAGGGGGCATCAGGCA 7288
QY	2543 CCCACAGGGGCTCCCACTCTCTCCCTGCCCACTGTGGCTTTCGATGTGTGCTCCAC 2602
DB	7287 CCCACAGGGGCTCCCACTCTCTCCCTGCCCACTGTGGCTTTCGATGTGTGCTCCAC 7228
QY	2603 GACCAACCGGCTCCCACTCTCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2651
DB	7227 GACCAACCGGCTCCCACTCTCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7168

PR	29-SEP-2000; 2000US-0236369P.
PR	29-SEP-2000; 2000US-0236370P.
PR	02-OCT-2000; 2000US-0236802P.
PR	02-OCT-2000; 2000US-0237037P.
PR	02-OCT-2000; 2000US-0237038P.
PR	02-OCT-2000; 2000US-0237039P.
PR	13-OCT-2000; 2000US-0237040P.
PR	13-OCT-2000; 2000US-0239935P.
PR	20-OCT-2000; 2000US-0240960P.
PR	20-OCT-2000; 2000US-0241785P.
PR	20-OCT-2000; 2000US-0241786P.
PR	20-OCT-2000; 2000US-0241787P.
PR	20-OCT-2000; 2000US-0241808P.
PR	20-OCT-2000; 2000US-0241826P.
PR	20-OCT-2000; 2000US-0242221P.
PR	01-NOV-2000; 2000US-0244617P.
PR	08-NOV-2000; 2000US-0246474P.
PR	08-NOV-2000; 2000US-0246475P.
PR	08-NOV-2000; 2000US-0246476P.
PR	08-NOV-2000; 2000US-0246477P.
PR	08-NOV-2000; 2000US-0246478P.
PR	08-NOV-2000; 2000US-0246523P.
PR	08-NOV-2000; 2000US-0246524P.
PR	08-NOV-2000; 2000US-0246525P.
PR	08-NOV-2000; 2000US-0246526P.
PR	08-NOV-2000; 2000US-0246527P.
PR	08-NOV-2000; 2000US-0246528P.
PR	08-NOV-2000; 2000US-0246532P.
PR	08-NOV-2000; 2000US-0246533P.
PR	08-NOV-2000; 2000US-0246534P.
PR	08-NOV-2000; 2000US-0246535P.
PR	08-NOV-2000; 2000US-0246536P.
PR	08-NOV-2000; 2000US-0246537P.
PR	08-NOV-2000; 2000US-0246538P.
PR	08-NOV-2000; 2000US-0246539P.
PR	08-NOV-2000; 2000US-0246540P.
PR	08-NOV-2000; 2000US-0246541P.
PR	08-NOV-2000; 2000US-0246542P.
PR	08-NOV-2000; 2000US-0246543P.
PR	08-NOV-2000; 2000US-0246544P.
PR	08-NOV-2000; 2000US-0246545P.
PR	08-NOV-2000; 2000US-0246546P.
PR	08-NOV-2000; 2000US-0246547P.
PR	08-NOV-2000; 2000US-0246548P.
PR	08-NOV-2000; 2000US-0246549P.
PR	08-NOV-2000; 2000US-0246550P.
PR	08-NOV-2000; 2000US-0246551P.
PR	08-NOV-2000; 2000US-0246552P.
PR	08-NOV-2000; 2000US-0246553P.
PR	08-NOV-2000; 2000US-0246554P.
PR	08-NOV-2000; 2000US-0246555P.
PR	08-NOV-2000; 2000US-0246556P.
PR	08-NOV-2000; 2000US-0246557P.
PR	08-NOV-2000; 2000US-0246558P.
PR	08-NOV-2000; 2000US-0246559P.
PR	08-NOV-2000; 2000US-0246560P.
PR	08-NOV-2000; 2000US-0246561P.
PR	08-NOV-2000; 2000US-0246562P.
PR	08-NOV-2000; 2000US-0246563P.
PR	08-NOV-2000; 2000US-0246564P.
PR	08-NOV-2000; 2000US-0246565P.
PR	08-NOV-2000; 2000US-0246566P.
PR	08-NOV-2000; 2000US-0246567P.
PR	08-NOV-2000; 2000US-0246568P.
PR	08-NOV-2000; 2000US-0246569P.
PR	08-NOV-2000; 2000US-0246570P.
PR	08-NOV-2000; 2000US-0246571P.
PR	08-NOV-2000; 2000US-0246572P.
PR	08-NOV-2000; 2000US-0246573P.
PR	08-NOV-2000; 2000US-0246574P.
PR	08-NOV-2000; 2000US-0246575P.
PR	08-NOV-2000; 2000US-0246576P.
PR	08-NOV-2000; 2000US-0246577P.
PR	08-NOV-2000; 2000US-0246578P.
PR	08-NOV-2000; 2000US-0246579P.
PR	08-NOV-2000; 2000US-0246580P.
PR	08-NOV-2000; 2000US-0246581P.
PR	08-NOV-2000; 2000US-0246582P.
PR	08-NOV-2000; 2000US-0246583P.
PR	08-NOV-2000; 2000US-0246584P.
PR	08-NOV-2000; 2000US-0246585P.
PR	08-NOV-2000; 2000US-0246586P.
PR	08-NOV-2000; 2000US-0246587P.
PR	08-NOV-2000; 2000US-0246588P.
PR	08-NOV-2000; 2000US-0246589P.
PR	08-NOV-2000; 2000US-0246590P.
PR	08-NOV-2000; 2000US-0246591P.
PR	08-NOV-2000; 2000US-0246592P.
PR	08-NOV-2000; 2000US-0246593P.
PR	08-NOV-2000; 2000US-0246594P.
PR	08-NOV-2000; 2000US-0246595P.
PR	08-NOV-2000; 2000US-0246596P.
PR	08-NOV-2000; 2000US-0246597P.
PR	08-NOV-2000; 2000US-0246598P.
PR	08-NOV-2000; 2000US-0246599P.
PR	08-NOV-2000; 2000US-0246600P.
PR	08-NOV-2000; 2000US-0246601P.
PR	08-NOV-2000; 2000US-0246602P.
PR	08-NOV-2000; 2000US-0246603P.
PR	08-NOV-2000; 2000US-0246604P.
PR	08-NOV-2000; 2000US-0246605P.
PR	08-NOV-2000; 2000US-0246606P.
PR	08-NOV-2000; 2000US-0246607P.
PR	08-NOV-2000; 2000US-0246608P.
PR	08-NOV-2000; 2000US-0246609P.
PR	08-NOV-2000; 2000US-0246610P.
PR	08-NOV-2000; 2000US-0246611P.
PR	08-NOV-2000; 2000US-0246612P.
PR	08-NOV-2000; 2000US-0246613P.
PR	08-NOV-2000; 2000US-0246614P.
PR	08-NOV-2000; 2000US-0246615P.
PR	08-NOV-2000; 2000US-0246616P.
PR	08-NOV-2000; 2000US-0246617P.
PR	08-NOV-2000; 2000US-0246618P.
PR	08-NOV-2000; 2000US-0246619P.
PR	08-NOV-2000; 2000US-0246620P.
PR	08-NOV-2000; 2000US-0246621P.
PR	08-NOV-2000; 2000US-0246622P.
PR	08-NOV-2000; 2000US-0246623P.
PR	08-NOV-2000; 2000US-0246624P.
PR	08-NOV-2000; 2000US-0246625P.
PR	08-NOV-2000; 2000US-0246626P.
PR	08-NOV-2000; 2000US-0246627P.
PR	08-NOV-2000; 2000US-0246628P.
PR	08-NOV-2000; 2000US-0246629P.
PR	08-NOV-2000; 2000US-0246630P.
PR	08-NOV-2000; 2000US-0246631P.
PR	08-NOV-2000; 2000US-0246632P.
PR	08-NOV-2000; 2000US-0246633P.
PR	08-NOV-2000; 2000US-0246634P.
PR	08-NOV-2000; 2000US-0246635P.
PR	08-NOV-2000; 2000US-0246636P.
PR	08-NOV-2000; 2000US-0246637P.
PR	08-NOV-2000; 2000US-0246638P.
PR	08-NOV-2000; 2000US-0246639P.
PR	08-NOV-2000; 2000US-0246640P.
PR	08-NOV-2000; 2000US-0246641P.
PR	08-NOV-2000; 2000US-0246642P.
PR	08-NOV-2000; 2000US-0246643P.
PR	08-NOV-2000; 2000US-0246644P.
PR	08-NOV-2000; 2000US-0246645P.
PR	08-NOV-2000; 2000US-0246646P.
PR	08-NOV-2000; 2000US-0246647P.
PR	08-NOV-2000; 2000US-0246648P.
PR	08-NOV-2000; 2000US-0246649P.
PR	08-NOV-2000; 2000US-0246650P.
PR	08-NOV-2000; 2000US-0246651P.
PR	08-NOV-2000; 2000US-0246652P.
PR	08-NOV-2000; 2000US-0246653P.
PR	08-NOV-2000; 2000US-0246654P.
PR	08-NOV-2000; 2000US-0246655P.
PR	08-NOV-2000; 2000US-0246656P.
PR	08-NOV-2000; 2000US-0246657P.
PR	08-NOV-2000; 2000US-0246658P.
PR	08-NOV-2000; 2000US-0246659P.
PR	08-NOV-2000; 2000US-0246660P.
PR	08-NOV-2000; 2000US-0246661P.
PR	08-NOV-2000; 2000US-0246662P.
PR	08-NOV-2000; 2000US-0246663P.
PR	08-NOV-2000; 2000US-0246664P.
PR	08-NOV-2000; 2000US-0246665P.
PR	08-NOV-2000; 2000US-0246666P.
PR	08-NOV-2000; 2000US-0246667P.
PR	08-NOV-2000; 2000US-0246668P.
PR	08-NOV-2000; 2000US-0246669P.
PR	08-NOV-2000; 2000US-0246670P.
PR	08-NOV-2000; 2000US-0246671P.
PR	08-NOV-2000; 2000US-0246672P.
PR	08-NOV-2000; 2000US-0246673P.
PR	08-NOV-2000; 2000US-0246674P.
PR	08-NOV-2000; 2000US-0246675P.
PR	08-NOV-2000; 2000US-0246676P.
PR	08-NOV-2000; 2000US-0246677P.
PR	08-NOV-2000; 2000US-0246678P.
PR	08-NOV-2000; 2000US-0246679P.
PR	08-NOV-2000; 2000US-0246680P.
PR	08-NOV-2000; 2000US-0246681P.
PR	08-NOV-2000; 2000US-0246682P.
PR	08-NOV-2000; 2000US-0246683P.
PR	08-NOV-2000; 2000US-0246684P.
PR	08-NOV-2000; 2000US-0246685P.
PR	08-NOV-2000; 2000US-0246686P.
PR	08-NOV-2000; 2000US-0246687P.
PR	08-NOV-2000; 2000US-0246688P.
PR	08-NOV-2000; 2000US-0246689P.
PR	08-NOV-2000; 2000US-0246690P.
PR	08-NOV-2000; 2000US-0246691P.
PR	08-NOV-2000; 2000US-0246692P.
PR	08-NOV-2000; 2000US-0246693P.
PR	08-NOV-2000; 2000US-0246694P.
PR	08-NOV-2000; 2000US-0246695P.
PR	08-NOV-2000; 2000US-0246696P.
PR	08-NOV-2000; 2000US-0246697P.
PR	08-NOV-2000; 2000US-0246698P.
PR	08-NOV-2000; 2000US-0246699P.
PR	08-NOV-2000; 2000US-0246700P.
PR	08-NOV-2000; 2000US-0246701P.
PR	08-NOV-2000; 2000US-0246702P.
PR	08-NOV-2000; 2000US-0246703P.
PR	08-NOV-2000; 2000US-0246704P.
PR	08-NOV-2000; 2000US-0246705P.
PR	08-NOV-2000; 2000US-0246706P.
PR	08-NOV-2000; 2000US-0246707P.
PR	08-NOV-2000; 2000US-0246708P.
PR	08-NOV-2000; 2000US-0246709P.
PR	08-NOV-2000; 2000US-0246710P.
PR	08-NOV-2000; 2000US-0246711P.
PR	08-NOV-2000; 2000US-0246712P.
PR	08-NOV-2000; 2000US-0246713P.
PR	08-NOV-2000; 2000US-0246714P.
PR	08-NOV-2000; 2000US-0246715P.
PR	08-NOV-2000; 2000US-0246716P.
PR	08-NOV-2000; 2000US-0246717P.
PR	08-NOV-2000; 2000US-0246718P.
PR	08-NOV-2000; 2000US-0246719P.
PR	08-NOV-2000; 2000US-0246720P.
PR	08-NOV-2000; 2000US-0246721P.
PR	08-NOV-2000; 2000US-0246722P.
PR	08-NOV-2000; 2000US-0246723P.
PR	08-NOV-2000; 2000US-0246724P.
PR	08-NOV-2000; 2000US-0246725P.
PR	08-NOV-2000; 2000US-0246726P.
PR	08-NOV-2000; 2000US-0246727P.
PR	08-NOV-2000; 2000US-0246728P.
PR	08-NOV-2000; 2000US-0246729P.
PR	08-NOV-2000; 2000US-0246730P.
PR	08-NOV-2000; 2000US-0246731P.
PR	08-NOV-2000; 2000US-0246732P.
PR	08-NOV-2000; 2000US-0246733P.
PR	08-NOV-2000; 2000US-0246734P.
PR	08-NOV-2000; 2000US-0246735P.
PR	08-NOV-2000; 2000US-0246736P.
PR	08-NOV-2000; 2000US-0246737P.
PR	08-NOV-2000; 2000US-0246738P.
PR	08-NOV-2000; 2000US-0246739P.
PR	08-NOV-2000; 2000US-0246740P.
PR	08-NOV-2000; 2000US-0246741P.
PR	08-NOV-2000; 2000US-0246742P.
PR	08-NOV-2000; 2000US-0246743P.
PR	08-NOV-2000; 2000US-0246744P.
PR	08-NOV-2000; 2000US-0246745P.
PR	08-NOV-2000; 2000US-0246746P.
PR	08-NOV-2000; 2000US-0246747P.
PR	08-NOV-2000; 2000US-0246748P.
PR	08-NOV-2000; 2000US-0246749P.
PR	08-NOV-2000; 2000US-0246750P.
PR	08-NOV-2000; 2000US-0246751P.
PR	08-NOV-2000; 2000US-0246752P.
PR	08-NOV-2000; 2000US-0246753P.
PR	08-NOV-2000; 2000US-0246754P.
PR	08-NOV-2000; 2000US-0246755P.
PR	08-NOV-2000; 2000US-0246756P.
PR	08

QY	2652	-----	2651
Db	7167	TAAGACACACGAGCAAAAGATCTTTGGGACATGGGTGGCTGACCTAGGATGCTTGG	7108
QY	2652	-----	2688
Db	7107	ATTGGATTATAGCCAGGCTTTTGTCTCCACAGATGCTCTGCTCTCTCCAGATTTTGAGC	7048
QY	2689	CTTTCATGTCGACACAGGTGTCGAGTTTCCTGCCAGACTGCGGAGCTGCGGGGTGAA	2748
Db	7047	CTCCATGTGCCACACAGGT - GTGCCAGTTTCTGCCAGACTGTGGGACAGCTGCGGGGTGAA	6989
QY	2749	GGGAGAGCTTGTGTGAACTGGATCCGAACCACTGCCAGAGCTACTGTTTCCCTTTGGAAC	2808
Db	6988	GGGAGAGGCTCTGCGGGACTGGATCAGAACCACCTGCCAGAGCTGCTATTTCCTTTGGAAC	6929
QY	2809	TGCCGGGCTGTGACACAGGAACCTGGAGGAGGCTGAGGACAGAGCTGGAGAGCTGCCATG	2868
Db	6928	TGCCGGGCTGTGACACAGAGCTGGAGGAGGCTGAGGATAGTCGGCTGGAGAGCTGCCATG	6869
QY	2869	TGTGGGCGCTGCATGCCAGAGAGGCTTGGAGGGGTTCGCCACTGCGGGACCCCAAGGGCCCT	2928
Db	6868	TGTGGGCGCTGCATGCCAGGAGAGGCTGGAGGGGTGCCAGTGGGGGCCCAACACGCTGT	6809
QY	2929	GGTGACAAAGGCTTTGGCTGTAGTTTATGCCCCCTTTGCCACTCACTACCCCAACACCTG	2988
Db	6808	-----	6753
QY	2989	GCTCGGCACATGAAGACTCAAGTGGTGAAGAACCCCTTCGCTGTGCCCTGTCCATAC	3048
Db	6752	GCCCGGCACATGAAGACACACAGTGGTGAAGAGCCCTTCGCTGTGCCCTGTCTCTAT	6693
QY	3049	GCCTCTGCTCATCTGGATAAACCCTGAACCGGACACAGCGGTGCCACACAGAGAAAAGCCC	3108
Db	6692	GCCTCTGCTCATCTGGATAAACCCTGAACCGGACACAGCGGTGCCATACAGAGAGAAAGCCC	6633
QY	3109	TACAAGTGCCTCTGTCTGCTATGCTGTGGCAACTGGCCCACTCAAGCGTCAATGT	3168
Db	6632	TACAAGTGCCTCTGTGCTCTTATGCTGTGGCAATCTTGGCCAACTCAAGCGTCAATGT	6573
QY	3169	CGCATCCACTCTGGTGAACAAACCTTTTCGGTGTAGCCCTTTGCAACTACAGCTGCACACAG	3228
Db	6572	CGCATCCACTCTGGTGAACAAACCTTTTCGGTGTAGCCCTTTGCAACTACAGCTGCACACAG	6513
QY	3229	AGTATGAACCTCAACGTCATGCTGGACACACGGGCGAGAGCCCTTCGGCTGTGCC	3288
Db	6512	AGCATGAACCTCAACGTCATGCTGGGACACACAGCGGAGAGCCCTTCGGCTGTGCC	6453
QY	3289	ACCTGGCGCTATACACAGGCGCACTGGGACAACTACAAAGCGTCTACAGAAGTGCATGGC	3348
Db	6452	ACCTGGCGCTATACCAAGGCGCACTGGGACAACTACAAAGCGTCTACAGAAGTGCATGGC	6393
QY	3349	CATGGTGAAGCAGGAGGCGCTGGTCTCTCTGCCCTGAGGGCTGGGCCCCACCTCATAGC	3408
Db	6392	CACGGTGGGGCAGGAGGCGCTGGTCTCTCTGCCCTGAGGGCTGGGCCCCACCTCATAGC	6333
QY	3409	CCACCTCTGTTTGTAGCACTCGGGGTCCAGCAGCCCTGGGTGCTACTGTGTAGCAGGGCT	3468
Db	6332	CCACCTCTGTTTGTAGCTCTCGGGGCCACACAGCCCTGGGGACTGCTGGCAGCGGGCT	6273
QY	3469	CTTCAATTCAGACTCACCTTGAACTAACTAGGTCTTTTACCTGGGGCTCTAGGAATTAGC	3528
Db	6272	GTCCACACAGACTCATCCTGAAC-----	6246
QY	3529	CCTATGCTCTGTGCAATTTTATACAAATGAATAGAAACCACTTTCCCTTTCTCCCCCGC	3588
Db	6245	TCCCTTCTTCCCCCATGTTTATACAGACGGAACAGNAGCACTT-----	6190
QY	3589	TGGTCAGGGGCTCCACACAGACTAACTTAGGCACTTATATGACACAGCTGTAATCCCATGG	3648
Db	6199	TGGCCAGGGCTCCACACAGACTAAGCTAGGCACTATAAGGACACAGCCCAACCCCATGGG	6130

Qy	3649	TCAGGGGG-CCATATAGACAGGGGA-CTTGTCTTACGTCACTACAGATGAGCTAAGT	3700
Db	6129	CGGGGGGGCCATATGAGACAGGGGACCTTGCCTTGACTGAGCACTTACGAGCTCAGT	6070
Qy	3707	GATTAGGGCTTCGATTACCGCCACTGCTCCAGAGGGCTATGATGAACCTGGTGGGAG	3766
Db	6069	GAGAGGGCCCTGTATTCACTCCATCTGCCCCAGGGGTGTGGAACAACCGGCTGGGG	6010
Qy	3767	-CTGCCAGCCTTTTAC- ---TGTTTTAACTTATTTCAGTGCCTTATATAAAGGAACA	3821
Db	6009	ACTGCCAGCCTCCACCTGTTTATTAACTTATTTCAGTGCCTTATATAAAGGAACA	5950
Qy	3822	CTAACAGA 3829	
Db	5949	CTAACAAA 5942	
RESULT 14			
ADA53558			
ID	ADA53558 standard; cDNA; 1456 BP.		
XX	ADA53558;		
XX	20-NOV-2003 (first entry)		
DT	Human coding sequence, SEQ ID 1126.		
DE	Cytostatic; Anti-inflammatory; Osteopathic; Neuroprotective; Nootropic;		
KW	Gene Therapy; human; secretory protein; membrane proteins; cancer;		
KW	inflammatory disease; osteoporosis; neurological disease; gene; ss.		
XX	Homo sapiens.		
OS	EPI293569-A2		
PN	19-MAR-2003.		
PD	21-MAR-2002; 2002EP-00006586.		
PF	14-SEP-2001; 2001JP-00328381.		
XX	24-JAN-2002; 2002US-0350435P.		
PR	(HELI-) HELIX RES INST.		
XX	(REAS-) RES ASSOC BIOTECHNOLOGY.		
PA	Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;		
XX	Yanamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Tamechika I;		
PI	Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;		
PI	WPI; 2003-395539/38.		
XX	P-PSDB; ADA55197.		
DR	New polynucleotides encoding full-length polypeptides, e.g. secretory		
XX	protein and/or membrane proteins, useful for developing medicines for diseases in		
CC	which the gene is involved, or as target molecules for gene therapy.		
PT	Claim 1; SEQ ID NO 1126; 205pp; English.		
XX	The present invention relates to novel human secretory or membrane		
CC	proteins (ADA54072-ADA55710) and their coding sequences (ADA52433-		
CC	ADA54071). The coding sequences are useful in the gene therapy of		
CC	diseases caused by abnormalities of the proteins, e.g. cancer,		
CC	inflammatory diseases, osteoporosis or neurological disease.		
XX	Sequence 1456 BP; 287 A; 495 C; 406 G; 268 T; 0 U; 0 Other;		
SQ	Query Match 29.0%; Score 1115.2; DB 7; Length 1456;		
	Best Local Similarity 87.4%; Pred. No. 2.8e-274;		
	Matches 1305; Conservative 0; Mismatches 144; Indels 44; Gaps 6;		
Qy	2341	CACCTACTCGAGCCACTGAAGCGGCACATGTGAGACACACAGCGGGAGAACCGCTTCGC	2400
Db	1	CACCTACTCGAGCCACTGAAGCGGCACATGTGAGACACACAGCGGGAGAACCGCTTCGC	60

XX	The invention relates to human nucleic acids (AAI57798-AAI61369) and the encoded polypeptides (AAI38642-AAI42213) with nootropic, immunosuppressant and cytostatic activity. The polynucleotides are useful in gene therapy. A composition containing a polypeptide or polynucleotide of the invention may be used to treat diseases of the peripheral nervous system, such as peripheral nervous injuries, peripheral neuropathy and localised neuropathies and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager Syndrome. Other uses include the utilisation of the activities such as: Immune system suppression, CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic CC and thrombolytic activity, cancer diagnosis and therapy, drug screening, CC assays for receptor activity, arthritis and inflammation, leukaemias and CC C.N.S disorders. Note: The sequence data for this patent did not form CC part of the printed specification									
XX	Sequence	1558 BP;	331 A;	456 C;	494 G;	277 T;	0 U;	0 Other;		
SQ	Query Match	21.2%;	Score	815.2;	DB	4;	Length	1558;		
	Best Local Similarity	85.7%;	Pred. No.	1.4e-197;						
	Matches	1004;	Conservative	0;	Mismatches	123;	Indels	45;	Gaps	7;
QY	2674	CTGCCAGACTGAGTCTTCATGTGCCACACAGGTGGTCCAGTTTCCTGCCAGACTGTGGG	2733							
Db	425	CCGACAGATTGAGCCCTCCATGTGCCACACAGGTGGTCCAGTTTCCTGCCAGACTGTGGG	484							
QY	2734	CAGCTGCGGGGTGAAGGGGAGAGCTTGTGTGAACTGGATCCGAACCACTGCCAGAGCTA	2793							
Db	485	CAGCTGCGGGGTGAAGGGGAGAGCTTGTGTGAACTGGATCCGAACCACTGCCAGAGCTG	544							
QY	2794	CTGTTCCCTTGGACCTCCGGGGCTGTGGACAGGAAGTGGAGAGGGTGGAGCAGCAGG	2853							
Db	545	CTATTCCCTTGGACCTCCGGGGCTGTGGACAGAGCTGGAGAGGGTGGAGGTAGTCGG	604							
QY	2854	CTGGGAGCTGCATGTGTGGCGCTGCATGCGAGAGAGGCTGGAGGGGTGGCACTGGG	2913							
Db	605	CTGGGAGCTGCCATGTGTGGCGCTGCATGCGAGAGAGGCTGGAGGGGTGGCACTGGG	664							
QY	2914	GGACCCAGGGCCCTGTGTGACAAAGGCTTGGCTGTGTAGTTTATGCCCTTTGGCACTCAC	2973							
Db	665	GGGCCCCAGGGCCCTGTGTGACAAAGGCTTGGCTGTGTAGCTCTGCCCTTTGCACTCAC	724							
QY	2974	TACCCCAACACCTGGGCTCGGCACATGAAGACTCACAGTGTGTGAGAAACCTTCCGCTGT	3033							
Db	725	TATCCCAACACCTGGGCTCGGCACATGAAGACACAGTGTGTGAGAAAGCCCTTCCGCTGC	784							
QY	3034	GCCCGCTGTCCATACGCTCTGTCTCATCTGGATAACCTGAAAGGGCACAGCGCGTCCAC	3093							
Db	785	GCCCGCTGTCTTATGCTCTGTCTCATCTGGATAACCTGAAAGGGCACAGCGCGTCCAT	844							
QY	3094	ACAGGAGAAAGCCCTACAGTGCCTCTGTCCGTATGCTGTGGCAACCTTGGCCCAAC	3153							
Db	845	ACAGGAGAAAGCCCTACAGTGCCTCTGTCCGTATGCTGTGGCAATCTTGCCCAAC	904							
QY	3154	CTCAAGGCTCATGTGTGCATCCATCTCTGGTGACAAACCTTTTCGGTGTAGCTTTGCAAC	3213							
Db	905	CTCAAGGCTCATGTGTGCATCCATCTCTGGTGACAAACCTTTTCGGTGTAGCTTTGCAAC	964							
QY	3214	TACAGTGCACACAGATATGAACCTTCAAGCTCATATGCTGCACACACGGGGGAGAAG	3273							
Db	965	TACAGTGCACACAGATATGAACCTTCAAGCTCATATGCTGCAGCACACAGGGGAGAAG	1024							
QY	3274	CCCTTCCGCTGTGCCACTGCGCTTATACACAGGCCACTGGGACAACTACAAGCGTCAAT	3333							
Db	1025	-CCTTCCGCTGTGCCACTGCGCTTATACACAGGCCACTGGGACAACTACAAGCGCCAC	1083							
QY	3334	CAGAAGGTGCATGGCCATGTGTGGAGAGAGGGCTGTCTCTCTGCCCTGTAGGGGTGG	3393							
Db	1084	CAGAAGGTGCATGGCCATGTGTGGAGAGAGGGCTGTCTCTCTGCCCTGTAGGGGTGG	1143							
QY	3394	GCCCAACCTCATAGCCACCTCTGTGTTTGTAGACTCGGGGTCCAGAGCCCTGGGTGCT	3453							

Db	1144	CCCCACCTCATATGCCACCCCTCTGTTTGTAGCTCTCGGGGGCCACCCAGCCCTGGGGA	1203							
QY	3454	ACTGTTAGCAGGGCTCTTTCATTTCAGACTCACCTTGAACCTAACTAGGTCTCTTTTACCTGGG	3513							
Db	1204	CTTGGCAGCCGGGTGTCCACACAGACTCATCTCTGAAC	1241							
QY	3514	GCTCTAGGAATTAGCCCTATATGCTTCCCTGCAATTTTATACAAATGAATAGAAACCACTTT	3573							
Db	1242	-----TAGGTCTCTTCTTCCCATGTTTATACAGACGCGACCAAGCCACCTT- 1289								
QY	3574	CCCTTCTCCCTCCCTGGTCAAGGGGCTCCACACAGACTAACTAGGCACTATATGGACCA	3633							
Db	1290	---TTTCTCCCTCCCTGGCTCCAGGGGCTCCACACAGACTTACGTAGGCACTATATAGGACCA	1346							
QY	3634	GCCTGAATCCCATGGTCAAGGGG-CCATATAGACACAGGGG-CTTGTCTTATAGTCAAGTA	3691							
Db	1347	GCCCAACCCCATGGCGGGGGGCCCATATGGACAGGGGACCTTGGCTTGACTGAGGCA	1406							
QY	3692	CCAGATGAGCTAAGTGATTAGGGGCTTGGATTACCCGCACTGCTCCAGAGGCTATGGA	3751							
Db	1407	CTTACAGAGCTCAGTGAGAGGGGCTTATTCACCTCCACTGCCCCCAGGGGCTGTGGA	1466							
QY	3752	TGAACCTGGTTGGAG-CTGCCCCAGGCTTTTAC---TGTTTTAACTTATTTTCACTGCTTT	3806							
Db	1467	CAAACCGGCTGGGGGACTGCCAGGCTCCACCTGTTTATTTAACTTATTTCACTGCTTT	1526							
QY	3807	ATAATAAAGGAAAACACTAACAGAAAAA	3838							
Db	1527	ATAATAAAGGAAAACACTAACAAAAA	1558							

Search completed: May 13, 2004, 23:07:59
Job time : 1399 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 13, 2004, 13:25:40 ; Search time 14620 Seconds
(without alignments)
11404.969 Million cell updates/sec

Title: US-09-673-994A-7
Perfect score: 3847
Sequence: 1 ggaggtgagatttattc.....gaaaaaaaaaaaaaaaaa 3847

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 3470272 seqs, 21671516995 residues
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:**

- 1: gb_ba:**
- 2: gb_hlg:**
- 3: gb_in:**
- 4: gb_om:**
- 5: gb_ov:**
- 6: gb_pat:**
- 7: gb_ph:**
- 8: gb_pl:**
- 9: gb_pr:**
- 10: gb_ro:**
- 11: gb_sts:**
- 12: gb_sy:**
- 13: gb_un:**
- 14: gb_vi:**
- 15: em_ba:**
- 16: em_fun:**
- 17: em_hum:**
- 18: em_in:**
- 19: em_mu:**
- 20: em_om:**
- 21: em_or:**
- 22: em_ov:**
- 23: em_pat:**
- 24: em_ph:**
- 25: em_pl:**
- 26: em_ro:**
- 27: em_scs:**
- 28: em_un:**
- 29: em_vi:**
- 30: em_htg_hum:**
- 31: em_htg_inv:**
- 32: em_htg_Other:**
- 33: em_htg_mus:**
- 34: em_htg_pln:**
- 35: em_htg_rod:**
- 36: em_htg_mam:**
- 37: em_htg_vrt:**
- 38: em_sy:**
- 39: em_htgo_hum:**
- 40: em_htgo_mus:**
- 41: em_htgo_Other:**

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	3836	99.7	3847	6	BD234712	BD234712 CNRE bind
2	2815.4	73.2	19296	2	AC114619	AC114619 Mus muscu
3	2628.4	68.3	2651	6	BD234722	BD234722 CNRE bind
4	2288	59.5	2289	6	BD234713	BD234713 CNRE bind
5	2092	54.4	274060	2	AC095184	AC095184 Rattus no
6	1988.4	51.7	2336	6	BD234721	BD234721 CNRE bind
7	1918.8	49.9	2220	10	BC058976	BC058976 Mus muscu
8	1518	39.5	2013	6	AX834458	AX834458 Sequence
9	1518	39.5	2013	9	AK096933	AK096933 Homo sapi
10	1505.8	39.1	2159	9	BC052282	BC052282 Homo sapi
11	1498	38.9	1975	6	AX274856	AX274856 Sequence
12	1367.6	35.5	1996	6	BD242880	BD242880 Secreter
13	1255.2	32.6	195880	9	AC074117	AC074117 Homo sapi
14	1115.4	29.0	1461	9	BC040650	BC040650 Homo sapi
15	1115.2	29.0	1456	6	AX714442	AX714442 Sequence
16	1115.2	29.0	1456	9	AK056765	AK056765 Homo sapi
17	960.4	25.0	112679	2	AC142219	AC142219 Rattus no
18	960.4	25.0	300078	2	AC109067	AC109067 Rattus no
19	835	21.7	1322	9	HSN805126	AL833946 Homo sapi
20	815.2	21.2	1558	6	AR338541	AR338541 Sequence
21	615.8	16.0	1824	6	AR083379	AR083379 Sequence
22	615.8	16.0	2069	10	MMU42383	U42383 Mus musculu
23	615.8	16.0	2181	10	BC009004	BC009004 Mus muscu
24	612.8	15.9	49743	2	AC114541	AC114541 Mus muscu
25	576.8	15.0	581	6	BD234719	BD234719 CNRE bind
26	559.6	14.5	49743	2	AC114541	AC114541 Mus muscu
27	528.4	13.7	851	6	BD082405	BD082405 87 human
28	526.8	13.7	2182	10	BC062083	BC062083 Rattus no
29	512.2	13.3	69843	2	AC025903	AC025903 Homo sapi
30	498	12.9	509	6	BD234714	BD234714 CNRE bind
31	485.8	12.6	500	6	BD234715	BD234715 CNRE bind
32	472.8	12.3	487	6	BD234716	BD234716 CNRE bind
33	450.2	11.7	586	6	BD234720	BD234720 CNRE bind
34	441.8	11.5	446	6	BD234717	BD234717 CNRE bind
35	399.2	10.4	496	6	AX400960	AX400960 Sequence
36	387.2	10.1	721	9	BC014386	BC014386 Homo sapi
37	387.2	10.1	1425	9	BC007361	BC007361 Homo sapi
38	385.6	10.0	1435	6	AX615059	AX615059 Sequence
39	385.6	10.0	2160	9	BC022061	BC022061 Homo sapi
40	385.6	10.0	2167	9	BC000057	BC000057 Homo sapi
41	385.6	10.0	2169	9	AK127593	AK127593 Homo sapi
42	383.6	10.0	706	6	BD021657	BD021657 Novel gen
43	383.6	10.0	706	6	BD101595	BD101595 Novel gen
44	380	9.9	809	6	AX053543	AX053543 Sequence
45	346.8	9.0	441	6	BD234718	BD234718 CNRE bind

ALIGNMENTS

RESULT 1
BD234712
LOCUS
CNRE binding factors and uses thereof.
DEFINITION
BD234712
ACCESSION
BD234712.1 GI:31044482
VERSION
JP 2002516664-A/6.
KEYWORDS
Mus musculus (house mouse)
SOURCE
Mus musculus
ORGANISM
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 3847)
AUTHORS
Chen, Y.B., Horiuchi, M., Dzu, V.J. and Tamura, K.
TITLE
CNRE binding factors and uses thereof
JOURNAL
Patent: JP 2002516664-A 6 11-JUN-2002;

BD234712 3847 bp DNA linear PAT 17-JUL-2003

Db 1741 GAGGGGACCTCTCAGGGTATGGGGCGGATGATCTGGGATCTTCGCTCTTACCAAGATA 1800
Qy 1801 TTAAGGAACCTGAGGGTCATCGAGTACGGGAAGTGCATGTCACAACAGCTGGCTCCCTTG 1860
Db 1801 TTAAGGAACCTGAGGGTCATCGAGTACGGGAAGTGCATGTCACAACAGCTGGCTCCCTTG 1860
Qy 1861 GTTCGGATTATGGGTACTGCTTGGGAGGAGATTCCACAAGACACCTCCCTCTTTAGTG 1920
Db 1861 GTTCGGATTATGGGTACTGCTTGGGAGGAGATTCCACAAGACACCTCCCTCTTTAGTG 1920
Qy 1921 GATACCTGAAGATTCCTTCGACGAAGTCTCTGGGGCCCTCGTGTGTGGAGACGATTTGCTA 1980
Db 1921 GATACCTGAAGATTCCTTCGACGAAGTCTCTGGGGCCCTCGTGTGTGGAGACGATTTGCTA 1980
Qy 1981 CTAGGCCAAGATCTGAGATTGGAAGAGGAGAGGAGAGGATGAAGGTACGGCCACAAC 2040
Db 1981 CTAGGCCAAGATCTGAGATTGGAAGAGGAGAGGAGAGGATGAAGGTACGGCCACAAC 2040
Qy 2041 GACACGCTCATGGGCTTTGAGAGAGACTCTCTGAAGGAGACTCTCAGGGGGCCAGACTGGA 2100
Db 2041 GACACGCTCATGGGCTTTGAGAGAGACTCTCTGAAGGAGACTCTCAGGGGGCCAGACTGGA 2100
Qy 2101 CTTCCCTATGGGCTGAGTACGACGAGTCTGGGGGGCGCGCGCACTAAAGTGGGAGGT 2160
Db 2101 CTTCCCTATGGGCTGAGTACGACGAGTCTGGGGGGCGCGCGCACTAAAGTGGGAGGT 2160
Qy 2161 GAAGTTGAGGAACACGACAGGGGTCCAGGGAGGGCCAGGGGTGAGAGGCCAGGCCAGCC 2220
Db 2161 GAAGTTGAGGAACACGACAGGGGTCCAGGGAGGGCCAGGGGTGAGAGGCCAGGCCAGCC 2220
Qy 2221 TGTACGCTGTGGGGGGCCGACAGGTGAGGGGCCGTGTGTGGGGGACAGAGGGGGGT 2280
Db 2221 TGTACGCTGTGGGGGGCCGACAGGTGAGGGGCCGTGTGTGGGGGACAGAGGGGGGT 2280
Qy 2281 GGGGGGGCCCGCTGCCCGACAGGTACTGCTACTGATCGCGGTGCGCTTCTGTGTC 2340
Db 2281 GGGGGGGCCCGCTGCCCGACAGGTACTGCTACTGATCGCGGTGCGCTTCTGTGTC 2340
Qy 2341 CACTACTCGAGCCACTGAAGCGGCAATGACAGACACACAGGGGGGAGAGCGTTCCGC 2400
Db 2341 CACTACTCGAGCCACTGAAGCGGCAATGACAGACACACAGGGGGGAGAGCGTTCCGC 2400
Qy 2401 TGTGGCGCTGCCATACGCTCAGCCAGTTCGTCTCACTGACGCGACATACCGCAC 2460
Db 2401 TGTGGCGCTGCCATACGCTCAGCCAGTTCGTCTCACTGACGCGACATACCGCAC 2460
Qy 2461 CATACTGGCGAGAGCCCTACCGTTGTCCCACTGCCCTTTGCTGCGACAGCCCTGGGC 2520
Db 2461 CATACTGGCGAGAGCCCTACCGTTGTCCCACTGCCCTTTGCTGCGACAGCCCTGGGC 2520
Qy 2521 AACCTGAGGGGGCATCAGCGCACCCACAGGGGCTCCCACTCTCTCCCTGCCCACTGT 2580
Db 2521 AACCTGAGGGGGCATCAGCGCACCCACAGGGGCTCCCACTCTCTCCCTGCCCACTGT 2580
Qy 2581 GGTCTTCGATGCTGTCTCCACGACCAACCGCGCTCCCACTCCACAGAGAGAGGGG 2640
Db 2581 GGTCTTCGATGCTGTCTCCACGACCAACCGCGCTCCCACTCCCAAGAGAGAGGGG 2640
Qy 2641 ACAATGCCCGACGATCAGAAAATGCGCTGATCCTGCGACAGCTTGATGTCTTATGTGCCA 2700
Db 2641 ACAATGCCCGACGATCAGAAAATGCGCTGATCCTGCGACAGCTTGATGTCTTATGTGCCA 2700
Qy 2701 CCAGGTGGTCCAGTTTCTCTGCCAGACTGTGGGAGCTGGGGGTGAAGGGGAGAGCTTG 2760
Db 2701 CCAGGTGGTCCAGTTTCTCTGCCAGACTGTGGGAGCTGGGGGTGAAGGGGAGAGCTTG 2760
Qy 2761 TGTGGAACTCGATCCGAAACCACTGCCAGAGCTACTGTTCCCTTGGACCTGCCGGGCTGT 2820
Db 2761 TGTGGAACTCGATCCGAAACCACTGCCAGAGCTACTGTTCCCTTGGACCTGCCGGGCTGT 2820
Qy 2821 GGACAGAACTGGAGGGGTGAGGGCAGCAGCTGGAGCTGCCATGTGTGGCGCTGC 2880

Db 2821 GGACAGAACTGGAGAGGGTGAAGGCGAGCAGGCTGGAGCTGCCATGTGTGGCGCTGC 2880
Qy 2881 ATGGAGAGAGGGTGAAGGGTTGCCACTTGGGGGACCCAGGGCCCTTGTGTGACAAAGGC 2940
Db 2881 ATGGAGAGAGGGTGAAGGGTTGCCACTTGGGGGACCCAGGGCCCTTGTGTGACAAAGGC 2940
Qy 2941 TTTGCCCTGTAGTTTATGCCCCCTTTGCCACTCACTTACCCCAACACCTTGGTTCGGCACATG 3000
Db 2941 TTTGCCCTGTAGTTTATGCCCCCTTTGCCACTCACTTACCCCAACACCTTGGTTCGGCACATG 3000
Qy 3001 AAGACTCACAGTGTGAGAAACCCCTTCGGTGTGCCGCTGTTCATACGGCTCTGTCTCAT 3060
Db 3001 AAGACTCACAGTGTGAGAAACCCCTTCGGTGTGCCGCTGTTCATACGGCTCTGTCTCAT 3060
Qy 3061 CTGATTAACCTGAAACCGCACCCAGCGCTCCACACAGAGAAAGCCCTCAAGTGGCCCC 3120
Db 3061 CTGATTAACCTGAAACCGCACCCAGCGCTCCACACAGAGAAAGCCCTCAAGTGGCCCC 3120
Qy 3121 CTCTGTCCGTATGCTGTGGCAACCTTGGCCAACTCAAGCGTCACTGTGTCATCCACTCT 3180
Db 3121 CTCTGTCCGTATGCTGTGGCAACCTTGGCCAACTCAAGCGTCACTGTGTCATCCACTCT 3180
Qy 3181 GGTGACAAACCTTTTCGGTGTAGCCTTTGCAACTACAGCTGCACACAGGTATGAACCTC 3240
Db 3181 GGTGACAAACCTTTTCGGTGTAGCCTTTGCAACTACAGCTGCACACAGGTATGAACCTC 3240
Qy 3241 AAAAGTCAATGCTGCGACACACCGGGGAGAGCCCTTCGCTGTGCCACCTGCGCCTAT 3300
Db 3241 AAAAGTCAATGCTGCGACACACCGGGGAGAGCCCTTCGCTGTGCCACCTGCGCCTAT 3300
Qy 3301 ACCACAGCCCATCGGACCACTACAAGGCTCATCAGAAGGTGATGCGCATGTGGAGCA 3360
Db 3301 ACCACAGCCCATCGGACCACTACAAGGCTCATCAGAAGGTGATGCGCATGTGGAGCA 3360
Qy 3361 GGAGGGCTGCTCTCTGCCCCCTGAGGGCTGGGGCCACCTCATAGCCACCCCTCTGTT 3420
Db 3361 GGAGGGCTGCTCTCTGCCCCCTGAGGGCTGGGGCCACCTCATAGCCACCCCTCTGTT 3420
Qy 3421 TTGAGCACTCGGGGTCCAGAGCCCTGCTACTGTGTAGCAGGGCTCTTCAATTCAGAC 3480
Db 3421 TTGAGCACTCGGGGTCCAGAGCCCTGCTACTGTGTAGCAGGGCTCTTCAATTCAGAC 3480
Qy 3481 TCACCTTGAACTAACTAGGTTCTTTTACCTGGGCTCTAGGAATTAGCCCTATGCCCTCT 3540
Db 3481 TCACCTTGAACTAACTAGGTTCTTTTACCTGGGCTCTAGGAATTAGCCCTATGCCCTCT 3540
Qy 3541 GCATTTTATACAAATGAACCTAGAAACCACTTTCCCTTTCTCCCGCTGGTCAAGGGCT 3600
Db 3541 GCATTTTATACAAATGAACCTAGAAACCACTTTCCCTTTCTCCCGCTGGTCAAGGGCT 3600
Qy 3601 CCACACAGACTAACTAGGCACTATATGGACCGCTGAATCCCATGGTCAAGGGGCCAT 3660
Db 3601 CCACACAGACTAACTAGGCACTATATGGACCGCTGAATCCCATGGTCAAGGGGCCAT 3660
Qy 3661 ATAGACAGGGGACTTGTCTTGTAGCTCAGTACACAGATGAGCTAAAGTATTAGGGCCCTTG 3720
Db 3661 ATAGACAGGGGACTTGTCTTGTAGCTCAGTACACAGATGAGCTAAAGTATTAGGGCCCTTG 3720
Qy 3721 ATTACCGCCACTGCTCCAGAGCTATGATGAATGTTGGTGGAGCTGCCAGCCCTTTT 3780
Db 3721 ATTACCGCCACTGCTCCAGAGCTATGATGAATGTTGGTGGAGCTGCCAGCCCTTTT 3780
Qy 3781 ACTGTTTTTAACTTATTTTCAAGTGTCTTATATAAAGGAAACACTAACAGAAAAA 3840
Db 3781 ACTGTTTTTAACTTATTTTCAAGTGTCTTATATAAAGGAAACACTAACAGAAAAA 3840
Qy 3841 AAAAAA 3847
Db 3841 AAAAAA 3847

RESULT 2
AC114619

```

LOCUS      AC114619                      199296 bp    DNA    linear    HTG 06-JUN-2002
DEFINITION Mus musculus clone RP24-86123, WORKING DRAFT SEQUENCE, 15 ordered
           pieces
ACCESSION  AC114619
VERSION    HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
KEYWORDS   Mus musculus (house mouse)
SOURCE     Mus musculus
ORGANISM   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE  1 (bases 1 to 199296)
AUTHORS    Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE      Mus musculus, clone RP24-86123
JOURNAL    Unpublished
REFERENCE  2 (bases 1 to 199296)
AUTHORS    Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
           Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L.,
           Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
           Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,
           Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
           Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
           Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
           Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
           Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
           Landers,T., Lehoczyk,J., Levine,R., Lindblad-Toh,K., Liu,G.,
           Maclean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
           McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,D.,
           Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,
           Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
           Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
           Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
           Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schuback,R.,
           Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
           Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
           Topham,K., Travers,M., Travis,N., Triggilio,J., Vassiliev,H.,
           Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
           Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
DIRECT SUBMISSION
TITLE      Submitted (10-MAR-2002) Whitehead Institute/MIT Center for Genome
JOURNAL    Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE  3 (bases 1 to 199296)
AUTHORS    Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
           Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L.,
           Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
           Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,
           Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
           Faro,S., Ferreira,P., FitzGerald,M., FitzHugh,W., Gage,D., Ginde,S.,
           Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L.,
           Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I.,
           Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K.,
           Lamazares,R., Landers,T., Lehoczyk,J., Levine,R., Lindblad-Toh,K.,
           Liu,G., Maclean,C., Macdonald,P., Major,J., Marquis,N.,
           Matthews,C., McCarthy,M., McEwan,P., McKernan,K., Meldrim,J.,
           Meneus,D., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
           Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P.,
           O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
           Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C.,
           Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S.,
           Schuback,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N.,
           Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S.,
           Theodore,J., Topham,K., Travers,M., Travis,N., Triggilio,J.,
           Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
           Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
DIRECT SUBMISSION
TITLE      Submitted (06-JUN-2002) Whitehead Institute/MIT Center for Genome
JOURNAL    Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT    On Jun 6, 2002 this sequence version replaced gi:19311148.
           All repeats were identified using RepeatMasker:
           http://ftp.genome.washington.edu/RN/RepeatMasker.html
           Smit, A.F.A. & Green, P. (1996-1997)
           ----- Genome Center
           Center: Whitehead Institute/ MIT Center for Genome Research
           Center code: WIBR

```

```

Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L24325
Center clone name: 861.23
----- Summary Statistics
Sequencing vector: Plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 193147 bases at least Q40
Consensus quality: 196185 bases at least Q30
Consensus quality: 197247 bases at least Q20
Insert size: 194000; agarose-fp
Insert size: 197896; sum-of-contigs
Quality coverage: 6.1 in Q20 bases; agarose-fp
Quality coverage: 6.0 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 15 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
*
1 728: contig of 728 bp in length
729 828: gap of 100 bp
829 2029: contig of 1201 bp in length
2129 2030: gap of 100 bp
2130 3803: contig of 1674 bp in length
3804 3903: gap of 100 bp
3904 6051: contig of 2148 bp in length
6052 6151: gap of 100 bp
6152 10082: contig of 3931 bp in length
10083 10182: gap of 100 bp
10183 18316: contig of 8134 bp in length
18317 18416: gap of 100 bp
18417 27114: contig of 8698 bp in length
27115 27214: gap of 100 bp
33934 33935: contig of 6720 bp in length
33935 34034: gap of 100 bp
34035 44466: contig of 10432 bp in length
44467 44567: gap of 100 bp
44567 59489: contig of 14923 bp in length
59490 59589: gap of 100 bp
59590 79430: contig of 19841 bp in length
79431 79530: gap of 100 bp
79531 99546: contig of 20015 bp in length
99546 123290: contig of 23645 bp in length
123291 123390: gap of 100 bp
123391 151687: contig of 28297 bp in length
151688 151787: gap of 100 bp
151788 159296: contig of 47509 bp in length.
159296 159296: Location/Qualifiers
1. 199296
   /organism="Mus musculus"
   /mol_type="genomic DNA"
   /db_xref="taxon:10090"
   /clone="RP24-86123"
   /clone_lib="RPCI-24 Male Mouse BAC"
1. 728
   /note="assembly_fragment"
829. 2029
   /note="assembly_fragment"
2130. 3803
   /note="assembly_fragment"
3904. 6051
   /note="assembly_fragment"
6152. 10082
   /note="assembly_fragment"
10183. 18316
   /note="assembly_fragment"

```

[illegible]

Db	184145	GTATTAAAGAACTGAGGTCATCGAGTACGGAGTGCAGTTCTCAACAGCTGGCTCC	184204
Qy	1858	TTGGTTCGGATTAATGGGTATCTGCTTGGGAGGGAGATTCCACAAGCACCTCCCTCTTTA	1917
Db	184205	TTGGTTCGGATTAATGGGTACTGCTTGGGAGGGAGATTCCACAAGCACCTCCCTCTTTA	184264
Qy	1918	GTGGATACTGAAGATTCTTCGACGAAGTCTCTGGGGCCCTGGTGTGGAGAGCGEATTTG	1977
Db	184265	GTGGATACTGAAGATTCTTCGACGAAGTCTCTGGGGCCCTGGTGTGGAGAGCGEATTTG	184324
Qy	1978	CTACTAGGCCAAGATCTGGAGTTTGAAGAGGAGAGAGAGATGAAGTGAAGCGCCAC	2037
Db	184325	CTACTAGGCCAAGATCTGGAGTTTGAAGAGGAGAGAGAGATGAAGTGAAGCGCCAC	184384
Qy	2038	AACGACAGCTCATGGGCTTTGAGAGAGACTCTGA	2072
Db	184385	AACGACAGCTCATGGGCTTTGAGAGAGACTCTGAAGGTGTGTTAAGGTCGTGGTTTA	184444
Qy	2073	-----	2072
Db	184445	GGATTAGTGGCAATTGGTAACTCAAGCCCTTTAGAAGCGGTCCAGCAGCCGAAACCC	184504
Qy	2073	-----	2072
Db	184505	ACTTCTACAACCTAGAACTCCCTAAAGTCCCTCAGCAGCATCATTTGCTGAGGTGAA	184564
Qy	2073	-----	2072
Db	184565	GTTTCGGGTTCCGGCTAAAGGTGTGTGAATACAGGGTGGGTAGCCCCAGTTCTTATGCG	184624
Qy	2073	-----	2072
Db	184625	CACCTCAAGAGGTTCAAGCAAGGAAATTATGTAATAATAGTACTCTTAAAACTCCTTTT	184684
Qy	2073	-----	2072
Db	184685	GACCAGTAGCAGCAGATCCTTTTGTACTGAGGTTTCAGTTTCTATCCTTAAGTCTCTCA	184744
Qy	2073	-----	2072
Db	184745	TTTATTTTGGGTAATTTTAAAGTCAGAGAAAACCATATAAAGCAATCTGCCCTTTC	184804
Qy	2073	-----	2072
Db	184805	TGTGTAAATTGCACACTATTGTTTCCACTTCTCTAAATTGTGCCTGCTTGAGMAGGCTT	184864
Qy	2073	-----	2072
Db	184865	TCAITTTCTAACCGCTTATATAAAGAGAGCCTTACTAAACAGAGCTTCAAAAAAG	184924
Qy	2073	-----	2072
Db	184925	ATCTTGGGATCTTTATATTTTTTCTGTCTCAAGNATTAAACGTTGCTTCAGTCAGCT	184984
Qy	2073	-----	2072
Db	184985	GCACCGTACTTCAAGCTCATGATATTTTCCATTTGTTAAAAATGACATGTCACCCGTGT	185044
Qy	2073	-----	2072
Db	185045	GTTAGTAATTAGGCTCAGAGGTAGGACCAATCAAGGATGAGGCAGGGTAAGTGGGGCTA	185104
Qy	2073	-----	2072
Db	185105	CATGTGAGAAGTGAGAGTTTGGAACTCAGTCTTAGCCATGTTGACCGGTTGTTTTTTT	185164
Qy	2073	-----AGGAGACTCTCAGGGGGCCAGACTGGACTTCCCTATGGGCTGAG	2117
Db	185165	CCCTTACGTTTCAACAGGAGACTCTCAGGGGGCCAGACTGGACTTCCCTATGGGCTGAG	185224
Qy	2118	TGACGACGAGTCTGGGGGGCGCGCACTAAGTGGGAGAGTGAAGTGAAGAAACGAC	2177

Db	185225	TGACGACGAGTCTGGGGGGGCGCCGCGCATACTAAGTCGGAGAGTGAAGTTGAGAACCCAGC	185284
Qy	2178	CAGGGGTCCAGGGGAGGCCAGGGGTGAGAGGCCAGGCCAGCCAGCTGTGACGTGTGTGGGG	2237
Db	185285	CAGGGGTCCAGGGGAGGCCAGGGGTGAGAGGCCAGGCCAGCCAGCTGTGACGTGTGTGGGG	185344
Qy	2238	GCACACAGGTGAGGGGCCGTGTTGTGGGCACAGAGGGCGGGT - GGGGGGCCCCCGCTGC	2296
Db	185345	GCACACAGGTGAGGGGGCGGTGTTGTGGGCAGAGGGCCGGGTGGGGGCCCCCGCTGC	185404
Qy	2297	CCCCACGGTTACTGTACTCATGCGGGCTGTGCGCTTTCTGTGCCACTACTCAGACCCACC	2356
Db	185405	CCCCTCGGCTACTGTACTCATGCGGGCTGTGCGCTTTCTGTGCCACTACTCAGACCCACC	185464
Qy	2357	TGAAGCGGCACATCCAGACACACAGCGGGGAGAGCGTTCCGCTGGCCGCTGCCCAT	2416
Db	185465	TGAAGCGGCACATCCAGACACACAGCGGGGAGAGCGTTCCGCTGGCCGCTGCCCAT	185524
Qy	2417	ACGNCTCAGCCCACTTCGTCAACTGACGGGACATACCCGACCCACTACTGCGAGAGAC	2476
Db	185525	ACGCTCAGCCCACTTCGTCAACTGACGGGACATACCCGACCCACTACTGCGAGAGAC	185584
Qy	2477	CCTACCGTTGCCCACTGCCCTTTGCTGCAGCAGCCTGGGCAACCTGAGGGGCATC	2536
Db	185585	CCTACCGTTGCCCACTGCCCTTTGCTGCAGCAGCCTGGGCAACCTGAGGGGCATC	185644
Qy	2537	AGCGCACCCACACAGGGGCTCCCACTCCTCCTCCGCCCAACCTGTGGGTTCGATGCTGTG	2596
Db	185645	AGCGCACCCACACAGGGGCTCCCACTCCTCCTCCGCCCAACCTGTGGGTTCGATGCTGTG	185704
Qy	2597	CTCCACGACCAACCCGGCTCCCACTGCTCCACAGAGCAGGAGGACAAATGCCCGACGAT	2656
Db	185705	CTCCACGACCAACCCGGCTCCCACTGCTCCACAGAGCAGGAGGACAAATGCCCGACGAT	185764
Qy	2657	CAGAAA-----	2662
Db	185765	CAGAAGGTAAAGAACACACAGGCATCTTGGGCATGGGTTGGTGTGATAGGAGTGCTT	185824
Qy	2663	-----ATGCGCTGATCCTGCCAGACTTGAGTC	2689
Db	185825	GGATTAATGGCCATATCTCATTTGCCACAGATGCGCTGATCCTGCGACACTTGATC	185884
Qy	2690	TTTCATGTGCCACCGTGGTGCAGTTTCCTGCCAGACTGTGGGCAGCTGCGGGGTGAAG	2749
Db	185885	TTTCATGTGCCACCGTGGTGCAGTTTCCTGCCAGACTGTGGGCAGCTGCGGGGTGAAG	185944
Qy	2750	GGGAGAGCTTGTGTGGAACTGGATCCGAACCACTGCCAGAGCTACTGTTCCTTTGGACCT	2809
Db	185945	GGGAGAGCTTGTGTGGAACTGGATCCGAACCACTGCCAGAGCTACTGTTCCTTTGGACCT	186004
Qy	2810	GCCGGGCTGTGGA CAGGAACTGGAGGAGGTGAGGCGACAGGCTGGAGCTGCATGT	2869
Db	186005	GCCGGGCTGTGGA CAGGAACTGGAGGAGGTGAGGCGACAGGCTGGAGCTGCATGT	186064
Qy	2870	GTGGGCGCTGCATGGCAGAGAGGCTGGAGGGTTGCCACTGGGGGACCCACAGGCCCTTG	2929
Db	186065	GTGGGCGCTGCATGGCAGAGAGGCTGGAGGGTTGCCACTGGGGGACCCACAGGCCCTTG	186124
Qy	2930	GTGACAAAGGCTTGCTGTAGTTTATGCCCTTTGCCACTCACTACCCACACCACTTGG	2989
Db	186125	GTGACAAAGGCTTGCTGTAGTTTATGCCCTTTGCCACTCACTACCCACACCACTTGG	186184
Qy	2990	CTCGGCACATGAAGACTCACAGTGTGAGAAACCTTCGCTGTGCCGCTGTCCATACG	3049
Db	186185	CTCGGCACATGAAGACTCACAGTGTGAGAAACCTTCGCTGTGCCGCTGTCCATACG	186244
Qy	3050	CCTCTGCTCATCTGATATACTGAAACGGCACAGGGGTCCACACAGAGAGAAAGCCCT	3109
Db	186245	CCTCTGCTCATCTGATATACTGAAACGGCACAGGGGTCCACACAGAGAGAAAGCCCT	186304
Qy	3110	ACAAGTGGCCCTCTGTCGATGCTGTGGCAAACCTGGCCCAACCTCAAGCGTCAATGTC	3169
Db	186305	ACAAGTGGCCCTCTGTCGATGCTGTGGCAAACCTGGCCCAACCTCAAGCGTCAATGTC	186364

2037 QY CAACGACGAGCTCATGGGCTTTGAGAGAGACTCTGAGGAGACTCTCAGGGGGCCAGACC 2096
841 Db CAACGACGAGCTCATGGGCTTTGAGAGAGACTCTGAGGAGACTCTCAGGGGGCCAGACC 900
2097 QY TGGACTTCCCTATGCGGTGAGTGACGAGAGTCTGGGGGGCGCGCGCACTAAAGTGCGGA 2156
901 Db TGGACTTCCCTATGCGGTGAGTGACGAGAGTCTGGGGGGCGCGCGCACTAAAGTGCGGA 960
2157 QY GAGTGAAGTTGAGGAACACGAGCCAGGGGTCACAGGGGAGGCCAGGGGTGAGAGGCCAGAGGCC 2216
961 Db GAGTGAAGTTGAGGAACACGAGCCAGGGGTCACAGGGGAGGCCAGGGGTGAGAGGCCAGAGGCC 1020
2217 QY AGCCTGTCTAGCTGTGTGGGGGCGCGAGAGTGAGGGGCGGTGTGTGGGGCAGAGAGCG 2276
1021 Db AGCCTGTCTAGCTGTGTGGGGGCGCGAGAGTGAGGGGCGGTGTGTGGGGCAGAGAGCG 1080
2277 QY GGGTGGGGGGCGCGCTGCCCTCAGCGTTACTGTACTCATGCTCGCGGTGTGGGTTTGGT 2336
1081 Db GGGTGGGGGGCGCGCTGCCCTCAGCGTTACTGTACTCATGCTCGCGGTGTGGGTTTGGT 1140
2337 QY GTCCCACTACTCGAGCCACTGAAGCGGCACATGACAGACACACAGCGGGGAGAGCGGTT 2396
1141 Db GTCCCACTACTCGAGCCACTGAAGCGGCACATGACAGACACACAGCGGGGAGAGCGGTT 1200
2397 QY CGCTGTGCGCGCTGCCCTAGCCATAGCCCTCAGCCAGTTCGCTCAACCTGACCGACATACCGG 2456
1201 Db CGCTGTGCGCGCTGCCCTAGCCCTAGCCCTCAGCCAGTTCGCTCAACCTGACCGACATACCGG 1260
2457 QY CACCCATATGCGGAGAGACCCCTACGTTGTGCCCACTGCCCTTTGCTGCGAGAGCGT 2516
1261 Db CACCCATATGCGGAGAGACCCCTACGTTGTGCCCACTGCCCTTTGCTGCGAGAGCGT 1320
2517 QY GGGGAACCTGAGCGGCGCATCGGACACCCACACAGAGGCTCCCATCTCCTCCCTGCCCAAC 2576
1321 Db GGGGAACCTGAGCGGCGCATCGGACACCCACACAGAGGCTCCCATCTCCTCCCTGCCCAAC 1380
2577 QY CTGTGGCTTTTCAGTGTGCTGCCAGCAACCCCGCTCCAGCTCCCAACAGAGCAGGA 2636
1381 Db CTGTGGCTTTTCAGTGTGCTGCCAGCAACCCCGCTCCAGCTCCCAACAGAGCAGGA 1440
2637 QY GGGGCAATGCGCCGACGATCAGAAATGCGGTGATCTGCGCAGACTTGAAGTCTTCATGT 2696
1441 Db GGGGCAATGCGCCGACGATCAGAAATGCGGTGATCTGCGCAGACTTGAAGTCTTCATGT 1500
2697 QY GCACCAAGTGTGCGAGTTTCTGCGACACTGCGGAGCTGCGGGGTGAAGGGGAGAG 2756
1501 Db GCACCAAGTGTGCGAGTTTCTGCGACACTGCGGAGCTGCGGGGTGAAGGGGAGAG 1560
2757 QY CTGTGTGAACTGGATCCGAACCACTGCGCAGAGTACTGTTCCTTGGAACTGCCGGG 2816
1561 Db CTGTGTGAACTGGATCCGAACCACTGCGCAGAGTACTGTTCCTTGGAACTGCCGGG 1620
2817 QY CTGTGACAGGAACTGGAGGGGTGAGGAGAGGCTGAGGAGCTGCCATGTGTGGGG 2876
1621 Db CTGTGACAGGAACTGGAGGGGTGAGGAGAGGCTGAGGAGCTGCCATGTGTGGGG 1680
2877 QY CTGCATGCGAGGAGAGGCTGAGGGGTTGCCACTGCGGAGACCCAGGGCCCTGTGTGACAA 2936
1681 Db CTGCATGCGAGGAGAGGCTGAGGGGTTGCCACTGCGGAGACCCAGGGCCCTGTGTGACAA 1740
2937 QY AGGCTTTGCTGTAGTTATGCGCTTTGCCACTACTACCCCAACCCAGCTGCGGCA 2996
1741 Db AGGCTTTGCTGTAGTTATGCGCTTTGCCACTACTACCCCAACCCAGCTGCGGCA 1800
2997 QY CATGAAGACTCAGTGTGTGAGAAACCCCTTCGCTGTGCCCTGTCCATAGCGCTCTGC 3056
1801 Db CATGAAGACTCAGTGTGTGAGAAACCCCTTCGCTGTGCCCTGTCCATAGCGCTCTGC 1860
3057 QY TCACTGTGATTAACCTGAAGCGGACAGCGGCTCCACAGAGAGAAAGCCCTTCAAGTG 3116
1861 Db TCACTGTGATTAACCTGAAGCGGACAGCGGCTCCACAGAGAGAAAGCCCTTCAAGTG 1920
3117 QY CCCCCTCTGCTGCTGTGGGCAACCTGGCCCACTCAAGCGCTCATGGTCCGATCCA 3176

1921 Db CCCCCTCTGCTGCTGCTGTGGCAACCTGGCAACCTCAAGCGCTCATGGTCCGATCCA 1980
3177 QY CTCTGTGTGACAAACCTTTTCGGTGTAGCCTTTTGCACACTACAGCTGCAACCCAGATATGAA 3236
1981 Db CTCTGTGTGACAAACCTTTTCGGTGTAGCCTTTTGCACACTACAGCTGCAACCCAGATATGAA 2040
3237 QY CTTCAAACGTCATATCTCTGCGACACACGCGGCGAGAGCCCTTCCGCTGTGCCACCTGCGC 3296
2041 Db CTTCAAACGTCATATCTCTGCGACACACGCGGCGAGAGCCCTTCCGCTGTGCCACCTGCGC 2100
3297 QY CTATACACAGGCGCACTGGGACAACTAAGCGTCTATCAGAGGTGCATGCGCATGGTGG 3356
2101 Db CTATACACAGGCGCACTGGGACAACTAAGCGTCTATCAGAGGTGCATGCGCATGGTGG 2160
3357 QY AGCAGAGGCGCTGTCTCTCTGCCCTCAGGGCTGGGCGCCACCTCATAGCCACCTC 3416
2161 Db AGCAGAGGCGCTGTCTCTCTGCCCTCAGGGCTGGGCGCCACCTCATAGCCACCTC 2220
3417 QY TGTGTTGAGCACTCGGGGTCCAGAGCCCTGGGTGTACTGTGAGCAGGGCTCTTCATTC 3476
2221 Db TGTGTTGAGCACTCGGGGTCCAGAGCCCTGGGTGTACTGTGAGCAGGGCTCTTCATTC 2280
3477 QY AGACTCACCTTTGAACCTAGGTCTTTTACCTGGGCTCTAGGAATTAGCCCTATGCC 3536
2281 Db AGACTCACCTTTGAACCTAGGTCTTTTACCTGGGCTCTAGGAATTAGCCCTATGCC 2340
3537 QY TCTGCACTTTTATCAAAATGAACCTAGAACCACTTTTCCCTTCTCCCGCTGGTCAGG 3596
2341 Db TCTGCACTTTTATCAAAATGAACCTAGAACCACTTTTCCCTTCTCCCGCTGGTCAGG 2400
3597 QY GGCTCCACACAGACTAACCTAGGCACTATATGACAGCCTGAATCCCATGGTCAGGGG 3656
2401 Db GGCTCCACACAGACTAACCTAGGCACTATATGACAGCCTGAATCCCATGGTCAGGGG 2460
3657 QY CCATATAGACAGGGGACTTGTCTTAGCTCACTACAGATGAGCTAAGTATTAGGGCC 3716
2461 Db CCATATAGACAGGGGACTTGTCTTAGCTCACTACAGATGAGCTAAGTATTAGGGCC 2520
3717 QY TTGGATTACCGGCACTGCTCCAGAGGCTATGGATGAAGTGGTGGAGCTGCCAGCC 3776
2521 Db TTGGATTACCGGCACTGCTCCAGAGGCTATGGATGAAGTGGTGGAGCTGCCAGCC 2580
3777 QY TTTTACTGTTTTAACTATTTCAGTGTCTTATTAAGAAACACTAAACAGAAAAAAA 3836
2581 Db TTTTACTGTTTTAACTATTTCAGTGTCTTATTAAGAAACACTAAACAGAAAAAAA 2640
3837 QY AAAAAAAAAA 3847
2641 Db AAAAAAAAAA 2651

RESULT 4

BD234713 2289 bp DNA linear PAT 17-JUL-2003
DEFINITION CNRE binding factors and uses thereof.
ACCESSION BD234713
VERSION BD234713.1 GI:33044483
KEYWORDS JP 2002516664-A/7.
SOURCE Mus musculus (house mouse).
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 2289)
AUTHORS Chen, Y. E., Horiuchi, M., Dzu, V. J. and Tamura, K.
TITLE CNRE binding factors and uses thereof
JOURNAL Patent: JP 2002516664-A 7 11-JUN-2002;
THE BRIGHAM AND WOMEN'S HOSPITAL INC
COMMENT OS Mus musculus (mouse)
PN JP 2002516664-A/7
PD 11-JUN-2002
PF 23-APR-1999 JP 2000545541
PR 24-APR-1998 US 60/082997

PI	YUQING E CHEN,MASATSUGU HORIUCHI,VICTOR J DZAU,KOICHI TAMURA	
PC	C12N15/09,A61K31/711,A61K38/00,A61K38/43,A61K45/00,A61K45/06,	
PC	A61K48/00,A61P9/04,A61P9/12,A61P43/00,C07K14/47,C12N1/15,C12N1/	
PC	19,C12N1/21,	
PC	C12N5/10,C12Q1/02,C12Q1/68,G01N33/15,G01N33/50,C12N15/00,PC	
PC	A61K37/02,	
PC	A61K37/48,C12N5/00	
CC	CDRE binding factors and uses thereof	
PH	Key Location/Qualifiers	
FT	source 1..2289	
FT	Location/Qualifiers	
FEATURES	Location/Qualifiers	
source	1..2289	
ORIGIN	Query Match 59.5%; Score 2288; DB 6; Length 2289;	
	Best Local Similarity 100.0%; Pred. No. 0;	
	Matches 2289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1198 ATGGAGGGGACGGGTACACACTTCTGGTGGCTGGTGGGTACGGTCCAGCTTCG 1257	
DB	1 ATGGAGGGGACGGGTACACACTTCTGGTGGCTGGTGGGTACGGTCCAGCTTCG 60	
QY	1258 ATTGGTTGGTACGAGCGATCAATCACTTCTGGTTCATGCTGCGATGCGCTCCGGT 1317	
DB	61 ATTGGTTGGTACGAGCGATCAATCACTTCTGGTTCATGCTGCGATGCGCTCCGGT 120	
QY	1318 TGTACGCCATGAATAGCATCCCGTGTCTGGTGGCTGGTGGGTACGGTCCAGCTTCG 1377	
DB	121 TGTACGCCATGAATAGCATCCCGTGTCTGGTGGCTGGTGGGTACGGTCCAGCTTCG 180	
QY	1378 ACCGAGGTACCACTGGTCTTCCGAGGGGGTACGGGGCTTCATGAATGGAAGCGCGCG 1437	
DB	181 ACCGAGGTACCACTGGTCTTCCGAGGGGGTACGGGGCTTCATGAATGGAAGCGCGCG 240	
QY	1438 GCGGCGGGAGCACTGAGCTGGATTCCGGGGCCGGGGAGGGGCTGCCAGGGCCCGC 1497	
DB	241 GCGGCGGGAGCACTGAGCTGGATTCCGGGGCCGGGGAGGGGCTGCCAGGGCCCGC 300	
QY	1498 ACCGTGTATGGGGGGCGTTCGTGGATCCTAAGAGCAAGACCGACGCGAGGGCCGAACCTG 1557	
DB	301 ACCGTGTATGGGGGGCGTTCGTGGATCCTAAGAGCAAGACCGACGCGAGGGCCGAACCTG 360	
QY	1558 GAGGGGCGGGGCGGGAGCGCTCCGGCCGGAGCGCGCTGGGCTGGAGCCGGTCAAGT 1617	
DB	361 GAGGGGCGGGGCGGGAGCGCTCCGGCCGGAGCGCGCTGGGCTGGAGCCGGTCAAGT 420	
QY	1618 GCCCGGAGGAGCAAGCCATCCAGCCCGTGAATGCGAGGGGTCAAAGGTCAAGG 1677	
DB	421 GCCCGGAGGAGCAAGCCATCCAGCCCGTGAATGCGAGGGGTCAAAGGTCAAGG 480	
QY	1678 GTCAGGGGCTTGAGCCGGGAGGAACAAGGGGTGGGGTCACTAGAGTGGGTCAAGTCAAG 1737	
DB	481 GTCAGGGGCTTGAGCCGGGAGGAACAAGGGGTGGGGTCACTAGAGTGGGTCAAGTCAAG 540	
QY	1738 GTGAGGGGAGCTCTCAGGGTTAGGGGCGGATGATCTGGATCTTCGCTCCTTACAGA 1797	
DB	541 GTGAGGGGAGCTCTCAGGGTTAGGGGCGGATGATCTGGATCTTCGCTCCTTACAGA 600	
QY	1798 GTATTAAAGGAACCTGAGGTCACTGAGTACGGGAAGTGCAGTTCAACAAGCTGGCTCC 1857	
DB	601 GTATTAAAGGAACCTGAGGTCACTGAGTACGGGAAGTGCAGTTCAACAAGCTGGCTCC 660	
QY	1858 TTGGTTCCGATTAATGGGTACTGCTTGGAGGGAGATTCCACAAGACCCCTCCCTCTTTA 1917	
DB	661 TTGGTTCCGATTAATGGGTACTGCTTGGAGGGAGATTCCACAAGACCCCTCCCTCTTTA 720	
QY	1918 GTGGATACGAAGATTCCTTCGAGGAAGTCTCGGGCCCTGGTGTGGAGAGCGAATTG 1977	

DB	721 GTGGATACCTAAGATTCTTCGACGAAGGTCTGGGGCCCTGGTGTGGAGAGCGATTG 780	
QY	1978 CTACTAGCCCAAGATCTGGAGTTTGAAGAGGAAGAGAGATGAAGTGTGACGGCCAC 2037	
DB	781 CTACTAGCCCAAGATCTGGAGTTTGAAGAGGAAGAGAGATGAAGTGTGACGGCCAC 840	
QY	2038 AAGACCACTCATGGGCTTTGAGAGAGACTCTGAAGAGACTCTCAGGGGGCCAGACCT 2097	
DB	841 AAGACCACTCATGGGCTTTGAGAGAGACTCTGAAGAGACTCTCAGGGGGCCAGACCT 900	
QY	2098 GGACTTCCCTATGGGTGAGTACGACGAGTCTGGGGGGCGGCGGCACTAAGTGGGGAG 2157	
DB	901 GGACTTCCCTATGGGTGAGTACGACGAGTCTGGGGGGCGGCGGCACTAAGTGGGGAG 960	
QY	2158 AGTGAAGTTGAGGAACAGCCAGGGGTTCAGGGGAGCCAGGGGTGAGAGCCAGGCCCA 2217	
DB	961 AGTGAAGTTGAGGAACAGCCAGGGGTTCAGGGGAGCCAGGGGTGAGAGCCAGGCCCA 1020	
QY	2218 GCCTGTGAGTGTGGGGGGCCGACAGGTGAGGGGCGGTGTGGGGGAGAGGGCGG 2277	
DB	1021 GCCTGTGAGTGTGGGGGGCCGACAGGTGAGGGGCGGTGTGGGGGAGAGGGCGG 1080	
QY	2278 GGTGGGGGGCCCGGCTGCCCCCAACGGTTACTGTACTCATGCCGGCTGTGGCTTTCGTG 2337	
DB	1081 GGTGGGGGGCCCGGCTGCCCCCAACGGTTACTGTACTCATGCCGGCTGTGGCTTTCGTG 1140	
QY	2338 TCCCACTACTCGAGCCACCTGAGCGGCAATGCGAGACACACAGCGGGGAGAGCCGCTC 2397	
DB	1141 TCCCACTACTCGAGCCACCTGAGCGGCAATGCGAGACACACAGCGGGGAGAGCCGCTC 1200	
QY	2398 CGCTGTGGCGCTGCCCATACGNTCAGCCAGTTCGTCAACTGACCGGCACTACCCGC 2457	
DB	1201 CGCTGTGGCGCTGCCCATACGNTCAGCCAGTTCGTCAACTGACCGGCACTACCCGC 1260	
QY	2458 ACCCATACTGGGAGAGCCCTACCGTTGCCCACTGCCCTTTGGCTGCGAGCAGCTG 2517	
DB	1261 ACCCATACTGGGAGAGCCCTACCGTTGCCCACTGCCCTTTGGCTGCGAGCAGCTG 1320	
QY	2518 GGCACACTGAGCGGCGCATCAGCGCACCCACACAGGGGCTTCCCACTCTCTCCCTGCCAAC 2577	
DB	1321 GGCACACTGAGCGGCGCATCAGCGCACCCACACAGGGGCTTCCCACTCTCTCCCTGCCAAC 1380	
QY	2578 TGTGGCTTCGATGCTGTCTCCAGCAACCCGGGCTCCAGTCCGACAGAGAGAG 2637	
DB	1381 TGTGGCTTCGATGCTGTCTCCAGCAACCCGGGCTCCAGTCCGACAGAGAGAG 1440	
QY	2638 GGGCAATGCCCCCGACGATCAGAAATGCGCTGATCTGCCAGACTTCAGTCTTCATGTG 2697	
DB	1441 GGGCAATGCCCCCGACGATCAGAAATGCGCTGATCTGCCAGACTTCAGTCTTCATGTG 1500	
QY	2698 CCACAGGTGGTGGCAGTTTCTCCAGACTGTGGGAGACTGTGGGAGTGCAGGGGAGAGAGC 2757	
DB	1501 CCACAGGTGGTGGCAGTTTCTCCAGACTGTGGGAGACTGTGGGAGTGCAGGGGAGAGAGC 1560	
QY	2758 TTGTGTGAATGGATCCGACCACTGCGAGACTTGTTCCTTGGACCTGCGGGGC 2817	
DB	1561 TTGTGTGAATGGATCCGACCACTGCGAGACTTGTTCCTTGGACCTGCGGGGC 1620	
QY	2818 TGTGACACAGGAATGAGAGGGGTGAGGGCAGAGGCTGGGAGTGCATGTGGGGGC 2877	
DB	1621 TGTGACACAGGAATGAGAGGGGTGAGGGCAGAGGCTGGGAGTGCATGTGGGGGC 1680	
QY	2878 TGCATGCGAGAGAGGCTGGAGGGGTTCGCACTGGGGAACCCAGGGCCCTGGTGACAAA 2937	
DB	1681 TGCATGCGAGAGAGGCTGGAGGGGTTCGCACTGGGGAACCCAGGGCCCTGGTGACAAA 1740	
QY	2938 GGCCTTGGCTGTAGTTTATGCCCTTTTCCCACTCACTACCCCAACCACTGGCTGGCAC 2997	
DB	1741 GGCCTTGGCTGTAGTTTATGCCCTTTTCCCACTCACTACCCCAACCACTGGCTGGCAC 1800	
QY	2998 ATGAAGACTCAAGTGTGAGAAAACCTTCCGCTGTGCGCTGTCCATACGCCCTCTGCT 3057	
DB	1801 ATGAAGACTCAAGTGTGAGAAAACCTTCCGCTGTGCGCTGTCCATACGCCCTCTGCT 1860	

* arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

* 1 270488: contig of 270488 bp in length
 * 270489: gap of unknown length
 * 270589: contig of 1405 bp in length
 * 271994: gap of unknown length
 * 272094: contig of 1967 bp in length.

FEATURES

Location/Qualifiers
 1..274060
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-9D7"
 1..1100
 /note="wgs_contig"
 214108..216519
 /note="wgs_contig"
 268481..270488
 /note="wgs_contig"

misc_feature

misc_feature

misc_feature

ORIGIN

Query Match 54.4%; Score 2092; DB 2; Length 274060;

Best Local Similarity 73.5%; Pred. No. 0;

Matches 3373; Conservative 0; Mismatches 315; Indels 902; Gaps 22;

Qy 124 CAGANAATTTTTTTTTCAGATGTGACGAGAGTTGGAGTAGTACCT--GANAACAT 181
 Db 265200 CCGACACATCTTCTGTGATGAGACGAGGCTGGGAATAGTGTAAATAGGACACAC 265141
 Qy 182 TAGTGACTTTGATAAAGNGNGGTTTTCATTTCCAGATGTAG-CCTGGGCTTC 240
 Db 265140 TAGAAGACCTGATTAAGAGATGAATCTTCTCGTTTCCCTGATACAGTCTTAGACTC 265081
 Qy 241 AGAAATNAC-----NGTTTTTTTTTGTGATAAATCGCCATTCTTTGTGCTCTT 296
 Db 265080 AGAAATTCAAAAATTTCTGTTTGTGATAATTTGTGGCCATCTCTCTCTGCCAC 265021
 Qy 297 CTGTGATTACAGGACAACTTCTGTGTTACTTT-----GGTCAAAC 342
 Db 265020 CATGTATCAGCAGGACAACTTCTATGGTACTCTGTGTTCCCGGAAGTGTCAA 264961
 Qy 343 CCATGGCTAACTGCTTAACTTTTGTAGCTGTGATCAGTCCGTGGCCAGACA 402
 Db 264960 CCGTGGCTAACTACCTCAACCTTTTGTAGCTGTGATCAGTCTGGCCCAAGACA 264901
 Qy 403 CTTCTGGGATGTACAGGGTGTGACAAATGACGTGATCATCTTTTCAAGCCCC 462
 Db 264900 CTTCTGGGATGTGACAGGGTGTGACAAATGACGTGATCATCTTTTCAAGCCCC 264841
 Qy 463 GAAACACAGTAGAGCTTCAGGACAGAGTGGCAAGAGGAACTGGAGGAGCACTGTCCA 522
 Db 264840 GAAACACAGTAGAGCTTCAGGACAGAGTGGCAAGAGGAACTGGAGGAGCACTGTCCA 264781
 Qy 523 CGGAGGGGCTGAGACACCGGCAAGTACAAAGAGAGGCAAGAGGATAGTGGT 582
 Db 264780 CGGAGGGGCTGAGAGATGGCAAGTACAAAGAGAGGCAAGAGGATAGTGGT 264721
 Qy 583 CAACCGGACCTGCCCCATGTGGACTTTTCTGAGCCCTTGGACCCGAGACTGATTTG 642
 Db 264720 CAACCGGACCTGCCCCATGTGGACTTTTCTGAGCCCTTGGACCCGAGACTGATTTG 264663
 Qy 643 TCCTTGTCTTTAGCTTACAGTGGGTATGAGTGTGAGGGGCTGGGTGCTTTCCT 702
 Db 264662 TCCTTGTCTTTAGCTTACAGTGGGTATGAGTGTGAGGGGCTGGGTGCTTTCCT 264603
 Qy 703 CAGCCCATTAAGAGGGGCCCCCACCCTCCAGCGGAGCTGGGAGGCTCTGCTG 762
 Db 264602 CAGCCCATTAAGAGGGTCTCT--CCCTCTGGACCGGACGCTGGGAGCTCTGCTG 264546
 Qy 763 TCCTCTTAAGCCCTCTTACT-CTCCTTGGGCTCATGCACTATCGGTTCTGTGCTGTGCT 821

Db 264545 TCCTCTCAGCCCTCCCTTACTCTCTTGGCTCATCGACTATCGGTTCTGTGCTGTGCT 264486
 Qy 822 CTGTGTGTGGAGGAGGAGCTGTGTCTGTATTTTACTCTGTGAAACACTTTATTA 881
 Db 264485 CTGTGTGTGGAGGAGGAGCTGTGTCTGTATTTTACTCTGTGAAACACTTTATTA 264426
 Qy 882 AGGACATCTTTTTTTTATTGGCGCTCTGTGACCCCTAGCCGCTTGACCCGCTCTCTGT 941
 Db 264425 AGGACATCTTTTTTTTATTGGCGCTCTGTGACCCCTAGCCGCTTGACCCGCTCTCTGT 264366
 Qy 942 GTACATCTTCAAGCAACACTTTTTCAGACTAAAGGCGCAAAACAAAGCTAATCTGTCTCAT 1001
 Db 264365 GTACATCTTCAAGCAACACTTTTTCAGACTAAAGGCGCAAAACAAAGCTAATCTGTCTCAT 264306
 Qy 1002 AGTGTCTATCTTTTACTCTCTACCTGGCCCATATGTGTGTGGCGCTGTGTCTGCATG 1061
 Db 264305 AGTGTCTATCTTTTCTCTCTACCTGGCCCATATGTGTGTGGCGCTGTGTCTGCATG 264246
 Qy 1062 CTTGTGTGTCTCCGAGCCCTCTGGAGAGGGGCGAGTGAATGTGGAGGAGGAGTGGAGT 1121
 Db 264245 --TGTGTGAGTCTCTGAGCCCTCTGGAGGGGAG-----CAGTGTGGATCAGGAGCTGGAAGT 264193
 Qy 1122 ----CTTAACTCTGAGTAAAGCAGCCGCTGGAGGCCATCCT--CGCTGTTCGCTTCAG 1176
 Db 264192 AGTGTAACTCTTAAAGCAGCCGCTGGAGGCCATCCTGTGGCTGTTCGCTTCAG 264133
 Qy 1177 GCAGCACTCTCTAGTCTGT--ATATGGAGGGGCGAGGTACCACTCTCTGTTGGCTGCT 1235
 Db 264132 GCATGACTCTCTAGTCTGT--ATATGGAGGGGCGAGGTACCACTCTCTGTTGGCTGCT 264073
 Qy 1236 TGGGTACGCTCAGCTCTGTGATTTGTTGGCTAGCAGCGCATCAATCACTTCTCTGCT 1295
 Db 264072 CGGCTCATTTTATTCAGCTTCTTATTTGCGGCTAGCAGCGCATCAATCACTTCTCTGCT 264013
 Qy 1296 CATGCTGAGCTGCGCTCGGCTTGTACAGCCCATGAACCTAGCATCCCTGTTCTCTGCT 1355
 Db 264012 CATGCTGCGCTGCGCTCCCG--TCTATATTCATAAATACGCTTCCCGAGTCTCTGCT 263954
 Qy 1356 GGTGCGGAAACGGAAGCGGTACGAGGTACAGCTGTGTTCTCGAGGGGGGTAGGGGG 1415
 Db 263953 GGTGCGGAAACGGAAGCGGTACGAGGTACAGCTGTGTTCTCGAGGGGGGTAGGGGG 263894
 Qy 1416 TCCATGATGAAGCGGCGCGCGGAGCGAGCTGAGCTGGATTCGCGGGGCC--GG 1474
 Db 263893 TCCATGATGAAGCGGCGCGCGGAGCGAGCTGAGCTGGATTCGCGGGGCCAGG 263834
 Qy 1475 GGCAGGGCTGCCCAAGGCCCGCAC--GTGTATGGGGCGGTTCTGTGATCTCTAAGAGCA 1533
 Db 263833 GGCAGGGCTGCCCAAGGCCCGCAC--GTGTATGGGGCGGTTCTGTGATCTCTAAGAGCA 263774
 Qy 1534 AGGACCGAGCGCGCGCGGAGCGCGGAGCGCTCGGAGCGCTCGGCGCGAGCG 1593
 Db 263773 AGGACCGAGCGCGCGGAGCGCGGAGCGCTCGGAGCGCTCGGCGCGAGCG 263714
 Qy 1594 CGTGGGCTGGAGCGCGCTCACGATGCCCGGAAGCAAGCAAGCCATCCACAGCCGTGA 1653
 Db 263713 CGTGGGCTGGAGCGCGCTCACGATGCCCGGAAGCAAGCAAGCCATCCACAGCCGTGA 263654
 Qy 1654 ATGCCAGGGGTCAAGGCTCAGGGCTCAGGGCTTGGAGCGGGAGCAAGGGGTGGGG 1713
 Db 263653 ATGCCAGGGGTCAAGGCTCAGGGCTCAGGGCTTGGAGCGGGAGCAAGGGGTGGGG 263594
 Qy 1714 TCAGTAGAGTGGGCTCAGGTCAGGGTGGAGGGGAGCTCCTCAGGGTTAGGGCGGATGAT 1773
 Db 263593 TCAGTAGAGTGGGCTCAGGTCAGGGTGAAGGGGTACTCTCAGGGTTAGGGCGGATGAT 263534
 Qy 1774 CTGGGATCTTCTGCTTACCGAGTATTAAGGACCTGAGGTCATCGAGTACGGAA 1833
 Db 263533 CTGGGATCTTCTGCTTACCGAGTATTAAGGACCTGAGGTCATCGAGTACGGAA 263474
 Qy 1834 GTGCAAGTTCACAAAGCTGGCTCTCTGTTGATTTATGGGTACTGCTTGGAGGAGAGAT 1893

QY 3202 AGCCTTTGCACTACAGCTCAACAGATGATGAACCTCAACGTCATATGCTGCGACAC 3261
 DB 261253 AGCCTTTGCACTACAGCTCAACAGATGATGAACCTCAACGTCATATGCTGCGACAC 261194
 QY 3262 ACGGGGAGAGAGCCCTTCCTGCTGCGACCTGCGGCTTATACACAGGCGCACTGGGACAC 3321
 DB 261193 ACGGGGAGAGAGCCCTTCCTGCTGCGACCTGCGGCTTATACACAGGCGCACTGGGACAC 261134
 QY 3322 TACAAGCGTCATCAGAAGAGTGCATGCGCATGCGGAGGAGGAGGCTGCTCTCTGCGC 3381
 DB 261133 TACAAGCGTCATCAGAAGAGTGCATGCGCATGCGGAGGAGGAGGCTGCTCTCTGCGC 261074
 QY 3382 CTTGAGGCTGGGCGCCACCTCATAGCCACCTCTCTGTTTGGAGCACTCGGGGTCAGCA 3441
 DB 261073 CTTGAGGCTGGGCGCCACCTCATAGCCACCTCTCTGTTTGGAGCACTCGGGGTCAGCC 261014
 QY 3442 GCCCTGGGTGCTACTGCTAGCAGGCTCTTCATTCAGACTCACCTTGAACCTAAGTGGTT 3501
 DB 261013 GCCCTGGGTGCTACTGCTAGCAGGCTCTTCACAGACTCACCTTGAACCTAAGTGGTT 260957
 QY 3502 CTTTACCTGGGCTCTAGGAATAGCCCTATGCTCTGCAATTTATACAAATGAACCTA 3561
 DB 260956 TTTTACCTGGGCTCTAGGAATAGCCCTATGCTCTGCAATTTATACAAATGAACCTA 260897
 QY 3562 G-AAACACACCTTTCCCTTTCTCCCGCGCTGCTCAGGGGCTCCACACAGACTAACTAGGC 3620
 DB 260896 GAAACACCTTTCCCTTTCTCCCGCGCTGCTCAGGGGCTCCACACAGACTAACTAGGC 260837
 QY 3621 ACTATATGACACCTGAAATCCATGCTGCTCAGGGGCTCATATAGACCA-GGGGACTTGT 3679
 DB 260836 ATTATATGACACCTGAAATCCATGCTGCTCAGGGGCTCATATAGACCA-GGGGACTTGT 260779
 QY 3680 TTACTCTACCTGACAGTGAAGTGAATGAGGCTTTGGATTCACCGCCACTGCTCCC 3739
 DB 260778 TTGCTCATGATGATGAGCTAAGTGAATGAGGCTTTGGATTCACCGCCACTGCTCCC 260719
 QY 3740 AGAGCTATGATGAAGTGTGGAGCTGCCAGCTTTTACTGTTTAACTATTTCOA 3799
 DB 260718 AGGGGCGACGATGATGTTGGAGCTGCCAGCTTTTACTGTTTAACTATTTCOA 260659
 QY 3800 GTGCTTTATATTAAGGAACACTTAACAGA 3829
 DB 260658 GTGCTTTATATTAAGGAACACTTAACAGA 260629

RESULT 6
 BD234721
 LOCUS 2336 bp DNA linear PAT 17-JUL-2003
 DEFINITION CNRE binding factors and uses thereof.
 ACCESSION BD234721
 VERSION BD234721.1 GI:33044491
 KEYWORDS JP 2002516664-A/15.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 Chien, Y. E., Horiuchi, M., Dzuu, V. J., and Tamura, K.
 1. (bases 1 to 2336)
 CNRE binding factors and uses thereof
 Patent: JP 2002516664-A 15 11-JUN-2002;
 THE BRIGHAM AND WOMEN'S HOSPITAL INC
 OS Mus musculus (mouse)
 PN JP 2002516664-A/15
 PD 11-JUN-2002
 PF 23-APR-1999 JP 2000545541
 PR 24-APR-1998 US 60/082997
 PI YUOLING E CHEN, MASATOSUGU HORIUCHI, VICTOR J DZAU, KOICHI TAMURA
 PC C12N15/09, A61K31/711, A61K38/00, A61K38/43, A61K45/00, A61K45/06,
 PC A61K48/00,
 PC A61P9/00, A61P9/04, A61P9/12, A61P43/00, C07K14/47, C12N1/15, C12N1/
 PC 19, C12N1/21,
 PC C12N5/10, C12Q1/02, C12Q1/68, G01N33/15, G01N33/50, C12N15/00, PC
 A61K37/02.

PC A61K37/48, C12N5/00
 CC CNRE binding factors and uses thereof
 FH key Location/Qualifiers
 FT source 1. .2336
 FT /organism="Mus musculus (mouse)"
 FT /organism="Mus musculus"
 FT /mol_type="genomic DNA"
 FT /db_xref="taxon:10090"

FEATURES
 source

ORIGIN

Query Match 51.7%; Score 1988.4; DB 6; Length 2336;
 Best Local Similarity 87.8%; Pred. No. 0;
 Matches 2328; Conservative 0; Mismatches 7; Indels 316; Gaps 2;

QY 1198 ATGGAGGGGGGAGGGGTACACCTCTCTGTTGGCTGGTTGGCTTACGCGTCAGCTTCTG 1257
 DB 1 ATGGAGGGGGGAGGGGTACACCTTCTCTGTTGGCTGGTTGGCTTACGCGTCAGCTTCTG 60
 QY 1258 ATTGGTTGGCTTAGCAGCGCATCAATCACTTCTCTGTTGGCTTACGCGTCAGCTTCTG 1317
 DB 61 ATTGGTTGGCTTAGCAGCGCATCAATCACTTCTCTGTTGGCTTACGCGTCAGCTTCTG 120
 QY 1318 TGTACAGCCATCAATACGATCCCGTGTCTCTGCGGTGGCGGAAGCGGAAGCGGGT 1377
 DB 121 TGTACAGCCATCAATACGATCCCGTGTCTCTGCGGTGGCGGAAGCGGAAGCGGGT 180
 QY 1378 ACGGAGTTACGAGCTGCTTTCGAGGGGGGTAGGGGGCTCATGAAATGGAAGCGGGC 1437
 DB 181 ACGGAGTTACGAGCTGCTTTCGAGGGGGGTAGGGGGCTCATGAAATGGAAGCGGGC 240
 QY 1438 GGGGGGGGAGGAGGAGGCTTCCGGGGCGGGG-CAGGGGCTGCCAGGGCGCG 1496
 DB 241 GGGGGGGGAGGAGGAGGCTTCCGGGGCGGGGCTGCCAGGGCGCG 300
 QY 1497 CACCTGTATGGGGGGGTTCGTGATCTTAAGAGCAAGACCGACGCGAGGGCGGAAT 1556
 DB 301 CACCTGTATGGGGGGGTTCGTGATCTTAAGAGCAAGACCGACGCGAGGGCGGAAT 360
 QY 1557 GGGAGGGGGGGGCGGGAGGAGGCTTCCGGGGCGGGGCTGGGGCTGGAGCGGTCACGA 1616
 DB 361 GGGAGGGGGGGGCGGGAGGAGGCTTCCGGGGCGGGGCTGGGGCTGGAGCGGTCACGA 420
 QY 1617 TGCCCCGAAGAGCAAGCCATCCACAGCCGTGAATTCGAGGGGTCAAAGTCAAG 1676
 DB 421 TGCCCCGAAGAGCAAGCCATCCACAGCCGTGAATTCGAGGGGTCAAAGTCAAG 480
 QY 1677 GGTACAGGGGCTTAGCCCGGAGAGCAAGGGGTGGGGTCAAGTGGGCTCAGGT 1736
 DB 481 GGTACAGGGGCTTAGCCCGGAGAGCAAGGGGTGGGGTCAAGTGGGCTCAGGT 540
 QY 1737 GGTGAGGGGGGACTCTCTCAGGGTTAGGGGGCGATGATCTGGGATCTTCTGCTTACAG 1796
 DB 541 GGTGAGGGGGGACTCTCTCAGGGTTAGGGGGCGATGATCTGGGATCTTCTGCTTACAG 600
 QY 1797 AGTATTAAAGAGCACTTACGGGTTCATCGAGTACGGGAAGTGCAGTTTCAACAGCTGGCT 1856
 DB 601 AGTATTAAAGAGCACTTACGGGTTCATCGAGTACGGGAAGTGCAGTTTCAACAGCTGGCT 660
 QY 1857 CTTGGTTCCGATTATGGGTACTGCTTGGAGGGAGATTCACAGCACTTCTCCCTCTTT 1916
 DB 661 CTTGGTTCCGATTATGGGTACTGCTTGGAGGGAGATTCACAGCACTTCTCCCTCTTT 720
 QY 1917 AGTGGATAGTGAAGATTCTCTCGAAGAGGCTCTCGGGGCTCTGGTGTGGAGAGCGATT 1976
 DB 721 AGTGGATAGTGAAGATTCTCTCGAAGAGGCTCTCGGGGCTCTGGTGTGGAGAGCGATT 780
 QY 1977 GCTACTAGCCCAAGATCTGGATTGAGAGCAAGAGAGAGGATGAGGTGAGCGCA 2036
 DB 781 GCTACTAGCCCAAGATCTGGATTGAGAGCAAGAGAGAGGATGAGGTGAGCGCA 840
 QY 2037 CAACGACCCAGCTCATGGGCTTTTGGAGAGACTCTTGAAGAGACTCTCAGGGGGCGGAC 2096

Db 841 CAAGCAGCAGCTCATGGCTTTTGAGAGAGACTCTGAAGGAGACTCTCAGGGGCCAGACC 900
QY 2097 TGGACTTCCCTATATGGCTGAGTACGACGAGTCTGGGGGGCGCGCGCACTAAGTGCAGA 2156
Db 901 TGGACTTCCCTATATGGCTGAGTACGACGAGTCTGGGGGGCGCGCGCACTAAGTGCAGA 960
QY 2157 GAGTGAAGTTGAGAAACAGCAGCGGTCCAGGGGAGCCAGGGGTGAGAGCCAGGCC 2216
Db 961 GAGTGAAGTTGAGAAACAGCAGCGGTCCAGGGGAGCCAGGGGTGAGAGCCAGGCC 1020
QY 2217 AGCTGTGAGTGTGGGGGGCGCAGGTGAGGGGCCGTGTGTGGGGCAGAGGGG 2276
Db 1021 AGCTGTGAGTGTGGGGGGCGCAGGTGAGGGGCCGTGTGTGGGGCAGAGGGG 1080
QY 2277 GGGTGGGGGGCCCCGCTGCCCCACCGTTACTGTACTCATGCGCGGTGTGCGCTTTTGGT 2336
Db 1081 GGGTGGGGGGCCCCGCTGCCCCACCGTTACTGTACTCATGCGCGGTGTGCGCTTTTGGT 1140
QY 2337 GTCCCACTACTCGAGCCACTGAGCGGCACATGCAGACACACAGCGGGGAGAGCGGTT 2396
Db 1141 GTCCCACTACTCGAGCCACTGAGCGGCACATGCAGACACACAGCGGGGAGAGCGGTT 1200
QY 2397 CCGTGTGGCGCGCTGCCCATACGNTCAGCCAGTTCGTCAACTGACGCGACATACCG 2456
Db 1201 CCGTGTGGCGCGCTGCCCATACGNTCAGCCAGTTCGTCAACTGACGCGACATACCG 1260
QY 2457 CACCCACTACTGGGAGAGCCCTACGTTGTCCCACTGCCCTTTGCGCTGACGACCT 2516
Db 1261 CACCCACTACTGGGAGAGCCCTACGTTGTCCCACTGCCCTTTGCGCTGACGACCT 1320
QY 2517 GGGCAACTGAGGGGCGCATCAGCGCACCCACACAGGGGCTCCCACTCTCCCTGCCCAAC 2576
Db 1321 GGGCAACTGAGGGGCGCATCAGCGCACCCACACAGGGGCTCCCACTCTCCCTGCCCAAC 1380
QY 2577 CTGTGGCTTTGATGCTGTGCTCAGACGACCGCGCTCCAGTCCAGCAGAGGAG 2636
Db 1381 CTGTGGCTTTGATGCTGTGCTCAGACGACCGCGCTCCAGTCCAGCAGAGGAG 1440
QY 2637 GGGGCAATGCCCGACCATCAGAAATGCGTGTATCTGCGCAGACTTGAGTCTTCATGT 2696
Db 1441 GGGGCAATGCCCGACCATCAGAAATGCGTGTATCTGCGCAGACTTGAGTCTTCATGT 1484
QY 2697 GCCACAGGTGTGTCAGTTTCTGCCAGACTGTGGGCGAGTGGGGGTGAAGGGGAG 2756
Db 1485 ----- 1484
QY 2757 CTGTGTGGAAGTGGATCCGACCACTGCCAGAGTACTGTTCCTTGGACCTGCCGGG 2816
Db 1485 ----- 1484
QY 2817 CTGTGGACAGAACTGGAGAGGGTGAGGGCAGCAGGCTGGGAGCTGCCATGTGTGGCG 2876
Db 1485 ----- 1484
QY 2877 CTGCATGAGGAGAGGCTGGAGGGGTGCCACTGGGGGACCCAGGCGCTGGTGACAA 2936
Db 1485 ----- 1484
QY 2937 AGGCTTGGCTAGTTATGTCCTTTGCGCCTTTGCCACTCACTACCCCAACCACTGGCTCGGCA 2996
Db 1485 -----A 1485
QY 2997 CATGAAGACTCAGTGTGAGAAACCTTCGCTGTGCGCGCTGCCATAGCGCTCTGC 3056
Db 1486 CATGAAGACTCAGTGTGAGAAACCTTCGCTGTGCGCGCTGCCATAGCGCTCTGC 1545
QY 3057 TCATCTGGATAACTGAAACGGCAGCGGTCCACAGAGAGAAAGCCCTACAAGTG 3116
Db 1546 TCATCTGGATAACTGAAACGGCAGCGGTCCACAGAGAGAAAGCCCTACAAGTG 1605
QY 3117 CCCCCTCTGTCGTGCTGTGGGCAACTGGCGCAACTCAGCGCTCATGTGCGATCCA 3176

Db 1606 CCCCCTCTGTCGGTATGCCTGGCAACCTGGCCCACTCAAGCGTCATGTCGCATCCA 1665
QY 3177 CTCTGGTGACAAACCTTTTCGGTGTAGCCTTTGCAACTACAGCTGCAACAGAGTATGAA 3236
Db 1666 CTCTGGTGACAAACCTTTTCGGTGTAGCCTTTGCAACTACAGCTGCAACAGAGTATGAA 1725
QY 3237 COTCAAAAGCTCATATGCTGCGACACAGGGGAGAGCCCTTCGCTGTGCAACCTGGC 3296
Db 1726 COTCAAAAGCTCATATGCTGCGACACAGGGGAGAGCCCTTCGCTGTGCAACCTGGC 1785
QY 3297 CTATACCAAGGCCACTGGGCAACTACAAAGCTCATCAGAAAGTGCATGCGCCATGTTGG 3356
Db 1786 CTATACCAAGGCCACTGGGCAACTACAAAGCTCATCAGAAAGTGCATGCGCCATGTTGG 1845
QY 3357 AGCAGAGGGCTGTGCTCTGTCGCCCTGAGGGCTGGGCCCACTCATAGCCACCCCTC 3416
Db 1846 AGCAGAGGGCTGTGCTCTGTCGCCCTGAGGGCTGGGCCCACTCATAGCCACCCCTC 1905
QY 3417 TGTCTTGAGCACTCGGGTCCAGCAGCCCTGGGTCTACTGCTAGCAGGGCTCTTTCATC 3476
Db 1906 TGTCTTGAGCACTCGGGTCCAGCAGCCCTGGGTCTACTGCTAGCAGGGCTCTTTCATC 1965
QY 3477 AGACTCACCTTGAACCTAGGTTCTTTTACCTGGGCTCTAGGAATTAGCCCTATGCC 3536
Db 1966 AGACTCACCTTGAACCTAGGTTCTTTTACCTGGGCTCTAGGAATTAGCCCTATGCC 2025
QY 3537 TCTCTGCAITTTATACAAATGAACCTAGAAACCACTTTTCTCTTCTCCCGCTGTGTCAGG 3596
Db 2026 TCTCTGCAITTTATACAAATGAACCTAGAAACCACTTTTCTCTTCTCCCGCTGTGTCAGG 2085
QY 3597 GGTCCACACAGCACTAACCTAGGCACTATATGGACAGCCCTGAATCCCATGCTCAGGGGG 3656
Db 2086 GGTCCACACAGCACTAACCTAGGCACTATATGGACAGCCCTGAATCCCATGCTCAGGGGG 2145
QY 3657 CCATATAGACACAGGGGACTTGTCTTAGCTCAGTACCAGATGAGCTAAGTATTAGGGCC 3716
Db 2146 CCATATAGACACAGGGGACTTGTCTTAGCTCAGTACCAGATGAGCTAAGTATTAGGGCC 2205
QY 3717 TTGATATCACCGCCACTGCTCCAGAGGCTATGATGAATGCTGTTGGAGCTGCCAGCC 3776
Db 2206 TTGATATCACCGCCACTGCTCCAGAGGCTATGATGAATGCTGTTGGAGCTGCCAGCC 2265
QY 3777 TTTTACTGTTTAACTTATTTCAGTCTTTTAAATAAGAAACACTAACAGAAAAA 3836
Db 2266 TTTTACTGTTTAACTTATTTCAGTCTTTTAAATAAGAAACACTAACAGAAAAA 2325
QY 3837 AAAAAAAAAA 3847
Db 2326 AAAAAAAAAA 2336

RESULT 7
BC058976
LOCUS
DEFINITION BC058976 2220 bp mRNA linear ROD 21-OCT-2003
IMAGE:5685033, complete cds.
ACCESSION BC058976
VERSION BC058976.1 GI:37589246
KEYWORDS MGC.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 2220)
Srausberg R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, P.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zebberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heish, F.,
Datchenko, M., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J.,

1146 CTGTGACAGGAACCTGGAGAGGGGTGAGGCGACGAGCTGGAGCTGCCATGTGTGGCG 1205
 2877 CTGCATGCGAGGAGAGGCTGGAGGGGTGCGACTGGGGGACCCAGGGCCCTGTGTGCAA 2936
 1206 CTGCATGCGAGGAGAGGCTGGAGGGGTGCGACTGGGGGACCCAGGGCCCTGTGTGCAA 1265
 2937 AGGCTTGGCTGTAGTTTATGCTCCCTTTGCTCCACTCACTACCCCAACCACTGGCTCGGCA 2996
 1266 AGGCTTGGCTGTAGTTTATGCTCCCTTTGCTCCACTCACTACCCCAACCACTGGCTCGGCA 1325
 2997 CATGAAGACTCAGAGGTGTGAGAAACCTTCCGCTGTGCTCCGCTGTCCATAGCCTCTGC 3056
 1326 CATGAAGACTCAGAGGTGTGAGAAACCTTCCGCTGTGCTCCGCTGTCCATAGCCTCTGC 1385
 3057 TCATCTGGATTAACCTGAAAGGCAACAGCGCTGCCACAGAGAGAAAGCCCTACAGTG 3116
 1386 TCATCTGGATTAACCTGAAAGGCAACAGCGCTGCCACAGAGAGAAAGCCCTACAGTG 1445
 3117 CCCCCTCTGTCGGTATGCTGTGGCAACCTGCGCAACCTCAAGCGTCAATGTGCAATCCA 3176
 1446 CCCCCTCTGTCGGTATGCTGTGGCAACCTGCGCAACCTCAAGCGTCAATGTGCAATCCA 1505
 3177 CTTCTGTGTGCAACCTTTTCGGGTGAGCGCTTTGCAACTACAGCTGCACACAGAGATGAA 3236
 1506 CTTCTGTGTGCAACCTTTTCGGGTGAGCGCTTTGCAACTACAGCTGCACACAGAGATGAA 1565
 3237 CTTCAAAAGCTCATATGCTGCGACACACGCGGAGAGCCCTTCCGCTGTGCCACTGCGC 3296
 1566 CTTCAAAAGCTCATATGCTGCGACACACGCGGAGAGCCCTTCCGCTGTGCCACTGCGC 1625
 3297 CTTATACACAGGCACTGGGACAACTTACAAAGCGTTCATCAGAGGTGCATGCCATGTGG 3356
 1626 CTTATACACAGGCACTGGGACAACTTACAAAGCGTTCATCAGAGGTGCATGCCATGTGG 1685
 3357 AGCAGAGGCGCTGTGCTCTGTCGCCCTGAGGCTGGGCGCCCACTCATAGCCACCCCTC 3416
 1686 AGCAGAGGCGCTGTGCTCTGTCGCCCTGAGGCTGGGCGCCCACTCATAGCCACCCCTC 1745
 3417 TGTTTTGAGCACTCGGGTCCAGCAGCCCTGGGTGCTACTGCTAGCAGGCTCTTCAATTC 3476
 1746 TGTTTTGAGCACTCGGGTCCAGCAGCCCTGGGTGCTACTGCTAGCAGGCTCTTCAATTC 1805
 3477 AGACTCACTTGAACCTAAGTAACTAGAACTAGTAACTAGTAACTAGTAACTAGTAACT 3536
 1806 AGACTCACTTGAACCTAAGTAACTAGAACTAGTAACTAGTAACTAGTAACTAGTAACT 1865
 3537 TCCTGCAATTTATACAAATGAACCTAGAACACCTTCCCTTTCTCCCGCTGTCAGG 3596
 1866 TCCTGCAATTTATACAAATGAACCTAGAACACCTTCCCTTTCTCCCGCTGTCAGG 1925
 3597 GGCTCCACACAGACTAACCTAGGCACTATATGACACAGCTGAATCCCATGCTCAGGGG 3656
 1926 GGCTCCACACAGACTAACCTAGGCACTATATGACACAGCTGAATCCCATGCTCAGGGG 1985
 3657 CCATATAGACCGGGGACTTGTCTAGCTCAGTACAGGATGAGTAACTAGTAACTAGTAACT 3716
 1986 CCATATAGACCGGGGACTTGTCTAGCTCAGTACAGGATGAGTAACTAGTAACTAGTAACT 2045
 3717 TTGGATTCACCGGCACTGTCTCCAGAGGCTATGATGAATGCTGTTGGAGCTGCCAGCC 3776
 2046 TTGGATTCACCGGCACTGTCTCCAGAGGCTATGATGAATGCTGTTGGAGCTGCCAGCC 2105
 3777 TTTTACTGTTTAACTTATTTTCACTGCTTTTATATTAAGGAACAACTAACAGAAAAAA 3836
 2106 TTTTACTGTTTAACTTATTTTCACTGCTTTTATATTAAGGAACAACTAACAGAAAAAA 2165
 3837 AAAAAAAAAA 3847
 2166 AAAAAAAAAA 2176

RESULT 8

AX834458
 LOCUS AX834458 2013 bp DNA linear PAT 15-DEC-2003
 DEFINITION Sequence 1582 from Patent EP1347046.
 ACCESSION AX834458
 VERSION AX834458.1 GI:39920593
 KEYWORDS Homo sapiens (human)
 SOURCE
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1
 Isogai,T., Sugiyama,T., Otsuki,T., Wakamatsu,A., Sato,H., Ishii,S.,
 Yamamoto,J.I., Isono,Y., Hio,Y., Otsuka,K., Nagai,K., Irie,R.,
 Tamechika,I., Seki,N., Yoshikawa,T., Otsuka,M., Nagahara,K. and
 Masuho,Y.
 Full-length cDNA sequences
 Patent: EP 1347046-A 1582 24-SEP-2003;
 Research Association for Biotechnology (JP)
 Location/Qualifiers
 source
 1..2013
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 ORIGIN
 Query Match 39.5%; Score 1518; DB 6; Length 2013;
 Best Local Similarity 88.6%; Fred. No. 0;
 Matches 1727; Conservative 0; Mismatches 179; Indels 44; Gaps 6;
 QY 1885 GAGGAGATTCACAAAGCAACCTCCCTCTTTAGTGGATATCTGAAGATTCCTTCGAGCAA 1944
 DB 102 GAGGAGATTCACAAAGCAACCTCCCTCTTTAGTGGATATCTGAAGACTCCCTCGAGCAA 161
 QY 1945 GGTCTGGGGCCCTGGTGTGGAGAGCAATTTGCTACTAGGCCAAGATCTGGAGTTTGA 2004
 DB 162 GGACCCGGGGCCCTGGTATTTGGAGAGTGAATTTGCTACTAGGCCAGGATCTGGAGTTTGA 221
 QY 2005 GAGGAAGAGAGAGAGATGAAGGTGACGGCCACACAGCAACAGCTCATGGCTTTTGAGAGA 2064
 DB 222 GAGGAAG 281
 QY 2065 GACTCTGAAGAGAGACTCTCAGGGGGCCAGACTGGAGCTTCCCTATGGCTGAGTGAGCAC 2124
 DB 282 GACTCTGAAGAGAGACTCTCAGGGGGCCAGAGCTGGGCTTCCCTATGGCTGAGTGAGCAT 341
 QY 2125 GAGTCTGGGGGGCGCGGCACTAAGTCGGAGAGTGAAGTTGAGGAACACAGCCAGGGGGT 2184
 DB 342 GAGTCTGGGGGGCGCGGCACTAAGTCGGAGAGTGAAGTTGAGGAACACAGCCAGGGGGT 401
 QY 2185 CCAGGGGAGGCCAGGGGTGAGAGGCCAGGCCAGCTGTGCTGCTGTGTGGGGGGCCGACA 2244
 DB 402 CCAGGGGAGGCCAGGGGTGAGAGGCCAGGCCAGGCCAGCTGTGCTGCTGTGTGGGGGGCCGACA 461
 QY 2245 GGTGAGGGGCCGCTGTTGTGGGGGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 2304
 DB 462 GGTGAGGGGCCGCTGTTGTGGGGGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 521
 QY 2305 TTACTGTACTCATGGCGCTGTGGCTTTTGTGTCCCACTACTCGAGCCACTGAGCGG 2364
 DB 522 CTACTGTACTCATGGCGCTGTGGCTTTTGTGTCCCACTACTCGAGCCACTGAGCGG 581
 QY 2365 CACATGACACACACAGGGGGGAGAGCGCTTCCGCTGTGGCGCTGCCCCATAGCCTCA 2424
 DB 582 CACATGACACACACAGGGGAGAGAGCGCTTCCGCTGTGGCGCTGCCCCATAGCCTCA 641
 QY 2425 GCCAGTTGTTCAACTGACCGGACATACCCGACCCGACCCGACCCGAGAGAGCCCTACCGT 2484
 DB 642 GCCAGTTGTTCAACTGACCGGACATACCCGACCCGACCCGAGAGAGCCCTACCGT 701
 QY 2485 TGTCCCACTGCCCCCTTTTGCCTGTCAGAGCCCTGGGCAACCTGAGCGGCGATCAGCGTACC 2544
 DB 702 TGTCCCACTGCCCCCTTTTGCCTGTCAGAGCCCTGGGCAACCTGAGCGGCGATCAGCGTACC 761

[illegible]

Db 1805 TAAGGACCGACCCCAACCCATGGCGGGGGGCCCATATGGAACAGGGGACCTTGCCCTTG 1865

Qy 3683 GCTCACGTACCATGAGTAACTGATTTAGGGCGCTTGGATTACCGCCACTGTCTCCAGA 3742

Db 1865 ACTGAGGCATCTCATGAGCTCACTGAGAGGGCCCTGTTATTCACTCCACTGCCCCCAGG 1924

Qy 3743 GGTATGAGTAACTGTTGGAG-CTGCCACGCTTTTAC- ---TGTTTTAACTTATTT 3797

Db 1925 GGCTGTGGACAAACCGGCTGGGAGCTGCCAGCTCCACCTGTGTTATTAACTTATTT 1984

Qy 3798 CAGTGCTTTATATAAGGAAACACTAAC 3826

Db 1985 CAGTGCTTTATATAAGGAAACACTAAC 2013

RESULT 9

AKO96933

LOCUS

DEFINITION Homo sapiens cDNA FLJ39614 fis, clone SMINT2000176, weakly similar to ZINC FINGER PROTEIN 83.

ACCESSION AKO96933

VERSION AKO96933.1 GI:21756542

KEYWORDS oligoapping; fis (full insert sequence).

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS Ninomiya, K., Wagatsuma, T., Kanda, K., Kondo, H., Yokoi, T., Kodaira, H., Furiya, T., Takahashi, M., Kikkawa, E., Omura, Y., Abe, K., Kanihara, K., Katsuta, N., Sato, K., Tanikawa, M., Yamazaki, M., Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S., Sugimoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Murakawa, K., Kanehori, K., Takahashi-Fujii, A., Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S., Nagahari, K., Masuho, Y., Nagai, K. and Isogai, T.

TITLE NEDO human cDNA sequencing project

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 2013)

AUTHORS Isogai, T. and Yamamoto, J.

TITLE Direct Submission

JOURNAL Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)

COMMENT NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.

FEATURES

source

1..2013

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="SMINT2000176"

/tissue_type="small intestine"

/clone_lib="SMINT2"

/note="cloning vector: pME18SFL3"

267..1706

/note="unnamed protein product"

/codon_start=1

/protein_id="BAC04903.1"

/db_xref="GI:21756543"

/translation="MGFERSDSGSLGARGPLPYGLSDDESGGRALSASFSEVEEPA
GFGEARGPACQLCGGPTGEGPCCGKAGGPGGLLPRLLYSCLRTYFVSHYSY
LKRHMTHSGEXPFRCGRCPYASQALVNLTRTRTHTGKPYKPCPCFACCSLHVS
RHQTHAGCTPTPCRCGCTPRPARPCEQEGAVPRPEADLLPDLSLHVF
GASSETPCGOLRGEGFHCGTSGSEPLPPLLFPWTCRGCGQLEEGEGRSLGAMCG

CDS

ORIGIN		CMRGAAGGASGGPQSPDKFACSLCPATHYPNHLARHMTKHSRKPFRCARCPYA SAHDLNLRKHQRVHTGEKPKYKPLCPYAGHNLNKRGRHISGDKPFRCSLONYSN QSNLKRHLRHTGEXFPFRCATCAVTTGHDNFKRQKQVYHGAGGPGLSASEGAP PHSPPSVLSSRGPALGTAGRAVHTDSS"
Query Match	39.5%; Score 1518; DB 9; Length 2013;	
Best Local Similarity	88.6%; Pred. No. 0;	
Matches 1727; Conservative	0; Mismatches 178; Indels 44; Gaps 6;	
QY	1885	GAGGAGATTCCCAAGCACCCTCCCTCTTTAGTGATCTGAAGATCTCTTCGACGAA 1944
DB	102	GAGGAGATTCCCAAGCACCCTCCCTCTTTAGTGATCTGAAGATCTCTTCGACGAA 161
QY	1945	GGTCTGGGCGCTGTGTGGAGAGCATTTGCTACTAGGCCAAGATCTGGAGTTTGA 2004
DB	162	GGACCCGGGCGCTGTGTGGAGAGCATTTGCTACTAGGCCAAGATCTGGAGTTTGA 221
QY	2005	GAGCAAGGAAGAGGATGAAGGTGACGGCCACACAGCAGCTCATGGGCTTTGAGAG 2064
DB	222	GAGGAAGAGGAGAGGAGGCGACGGCAACGTGACCATGTGGCTTCGAGAGA 281
QY	2065	GACTCTGAAGAGACTCTCAAGGGGCGCAGACCTTGGAATTCCCTATGGGCTGAGTGACGAC 2124
DB	282	GACTCGGAAGAGACTCTCTGGGGCGCAGGCTTGGGCTTCCCTATGGGCTGAGCGACGAT 341
QY	2125	GAGTCTGGGGCGCGCGCACTAAGTGGGAGAGTGAAGTTGAGGAACAGCAGCGGGGT 2184
DB	342	GAGTCTGGGGCGCGCGGCACTAAGTGGGAGAGTGAAGTTGAGGAGCAGCAGCGGGGT 401
QY	2185	CCAGGGAGGCGCAGGGGTGAGAGGCCAGGGCCAGCCCTGTGAGTGTGTGGGGGCGGACA 2244
DB	402	CCAGGGAGGCGCAGGGGTGAGAGGCCAGGGCCAGCCCTGTGAGTGTGTGGGGGCGGACA 461
QY	2245	GGTGAGGGCGCGTGTGTGGGGCAGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 2304
DB	462	GGTGAAGGGCGGTGTGTGGGGCAGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 521
QY	2305	TTACTGTACTCATGCGGCTGTGGGCTTTCGTGTCCCACTACTGAGCCACCTGAAGCGG 2364
DB	522	CTACTGTACTCATGCGGCTGTGACCTTTCGTGTCCCACTACTGAGCCACCTGAAGCGG 581
QY	2365	CACATGACACACAGGGGGGAGAGCCCTTCGCTGTGGCGCTGCCATCAGGNTCA 2424
DB	582	CACATGACACACAGGGGGGAGAGCCCTTCGCTGTGGCGCTGCCATCAGGNTCA 641
QY	2425	GCCAGTTTCTCAACCTGACCGACATACCCGACCCCACTATGGCGAGAGCCCTACCGT 2484
DB	642	GCCAGTTTCTCAACCTGACCGACATACCCGACCCCACTATGGCGAGAGCCCTACCGG 701
QY	2485	TGTCCCACTGCGCCCTTTGCTGTGAGAGCCCTGGGCAACCTGAGCGGCGATCAGCGCAC 2544
DB	702	TGTCCCACTGCGCCCTTTGCTGTGAGAGCCCTGGGCAACCTGAGCGGCGATCAGCGTACC 761
QY	2545	CACACAGGGCGCTCCACCTCCCTCCCTCCCACTGTGTGCTTTTCCATGTGTGTCCACGA 2604
DB	762	CACGAGGGCGCCCACTCTCTCCCTCCCTCCCACTGTGTGCTTTTCCATGTGTGTCCACGA 821
QY	2605	CCAAACCGGCGCTCCAGTCCCAACAGAGAGAGGGGGAACAATGCCCGGACGATCAGAAAT 2664
DB	822	CCAGCCCGGCGCTCCAGTCCCAACAGAGAGAGGGGGGGGGGGGGGGGGGGGGGGGGG 881
QY	2665	GGCTGTATCTCCAGACTTGTGATCTTCACTGTGCCACAGGTGTGGCTTCTCTGCCA 2724
DB	882	GCTGTGCTCTTCCAGATTGTAGCTTCAATGTGCCACAGGTGTGGCTTCTCTGCCA 941
QY	2725	GACTGTGGGCGAGCTGGGGGTGAAGGGGAGAGCTTTGTGTGGAATCTGGATCCGAACCACTG 2784
DB	942	GACTGTGGGCGAGCTGGGGGTGAAGGGGAGAGCTTTGTGTGGAATCTGGATCCGAACCACTG 1001
QY	2785	CCAGAGCTACTGTTCCCTTGGACCTGCGGGGCTGTGACAGGACTGGAGGGGTGAG 2844
DB	1002	CCAGAGCTGCTATTCCCTTGGACCTGCGGGGCTGTGACAGGACTGGAGGGGTGAG 1061

QY	2845	GGCAGCAGGCTGGAGCTGCCATGTGTGGGGGCTGCATGGAGAGAGGCTGGAGGGGTT 2904
DB	1062	GCTAGTGGGCTGGAGCTGCCATGTGTGGGGGCTGCATGGAGAGAGGCTGGAGGGGTT 1121
QY	2905	GCCACTGGGGGACCCCAAGGGCCCTGGTGACAAAGGCTTTGCCTGTAGCTTTATGCCCTTT 2964
DB	1122	CCAGTGGGGGCCCCAGGGCCCCAGTGACAAAGGCTTTGCCTGTAGCTTTATGCCCTTT 1181
QY	2965	GCCACTCACTACCCCAACCCTGGCTGGGACATGAGACTCACAGTGTGAGAAACCC 3024
DB	1182	GCCACTCACTATCCCAACCCTGGCTGGGACATGAGACTCACAGTGTGAGAAACCC 1241
QY	3025	TTCCGCTGTGCGGTGTCATACGCTCTGCTCATCTGGATAACCTGAAACGGCACCAG 3084
DB	1242	TTCCGCTGTGCGGTGTCATACGCTCTGCTCATCTGGATAACCTGAAACGGCACCAG 1301
QY	3085	GGCTGCCACACAGGAGAAAGCCCTAAGTGGCCCTCTGTCCGTATGCTGTGGCAAC 3144
DB	1302	CGCGTCCATACAGGAGAGAGCCCTAAGTGGCCCTCTGTCCCTTATGCTGTGGCAAT 1361
QY	3145	CTGGCCAAACCTCAAGCTGTCATGCTGGTGCATCACTCTGGTGACAAACCTTTTCGGTGTAGC 3204
DB	1362	CTGGCCAAACCTCAAGCTGTCATGCTGGTGCATCACTCTGGTGACAAACCTTTTCGGTGTAGC 1421
QY	3205	CTTTGCAACTACAGCTGCCAACGAGATATGAACCTCAAACTCATATGCTGCGACACAG 3264
DB	1422	CTTTGCAACTACAGCTGCCAACGAGATATGAACCTCAAACTCATATGCTGCGGACACACA 1481
QY	3265	GGGAGAGAGCCCTTCGCTGTGGCCACCTGCGCTATACACAGGCGCTGGGACAACTAC 3324
DB	1482	GGGAGAGAGCCCTTCGCTGTGGCCACCTGCGCTATACACAGGCGCTGGGACAACTAC 1541
QY	3325	AAGCGCTCATAGAGGTCATGGCCATGTGGAGCAGAGGCGCTGTCTCTGCGCCCT 3384
DB	1542	AAGCGCCACCAAGGTCATGGCCACGTCAGGCGAGGCGCTGTCTCTGCGCTCT 1601
QY	3385	GAGGCTGGGCGCCCACTCATAGCCACCTCTGTTTGGACCTCGGGGTCAGAGCC 3444
DB	1602	GAGGCTGGGCGCCCACTCATAGCCACCTCTGTTTGGACCTCGGGGTCAGAGCC 1661
QY	3445	CTGGGTGTCTACTGTGTAGAGGGCTCTTCACTCAGACTCACCTTGAACCTAGGTCTTT 3504
DB	1662	CTGGGACTGTGTGGCAGCGGGCTGTCCACACAGACTCATCTGAAC----- 1708
QY	3505	TTACTGGGGCTCTAGGAATTAGCCCTATGCTCTGCTATTTTATACAAATGAACAGAA 3564
DB	1709	-----TAGGTCCTTCTTCCCATGTTTTATACAGAGCGACCAAGAA 1748
QY	3565	ACCACCTTTTCCCTTTCTCCCGCTGGTCAGGGGCTCCACACAGACTAACCTTAGGCACCTA 3624
DB	1749	GCCACCTT-----TTTCTCCCGCTGGCCAGGGGCTCCACACAGACTAACCTTAGGCACCTA 1804
QY	3625	TATGGACAGCTGAATCCATGTGTGAGGGG-CCATATAGACAGGGGA-CTTGTCTTA 3682
DB	1805	TAAGGACAGCCCAACCCCATGGGGGGGGGGCCCATATGGACAGGGGACCTTCCCTTG 1864
QY	3683	GCTCAGTACACAGATGAGCTAAGTATGAGGCTTTGGAATCACCCGCACTCTCCACAGA 3742
DB	1865	ACTGAGGCTCTCATGAGCTCAGTGAGAGGGGCGCTGTATTCACTCCACTGCCCCAGG 1924
QY	3743	GGCTATGAGTGAACCTGGTTGGAG-CTGCCAGCGCTTTTAC-----TGTTTTAACTTATTT 3797
DB	1925	GGCTGTGGACAAACCGGCTGGGGGACTGCCAGCGCTCCACCTGTTTATTTAACTTATTT 1984
QY	3798	CAGTGTCTTATATTAAGGAACACTAAC 3826
DB	1985	CAGTGTCTTATATTAAGGAACACTAAC 2013

RESULT 10
BC052282
LOCUS

BC052282 2159 bp mRNA linear PRI 08-OCT-2003

DEFINITION Homo sapiens hypothetical protein FLJ32203, mRNA (cDNA clone MGC:59968 IMAGE:5467936), complete cds.

ACCESSION BC052282

VERSION BC052282.1

KEYWORDS GI:30354479

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 2159)

AUTHORS Strausberg, R.L., Peingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Schenker, C.M., Schuler, G.D., Altshuler, S.F., Zeeberg, K.H., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heien, F., Diatchenko, L., Marudina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stachenko, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J., McKernan, R.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickinson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, J., Smallus, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.

TITLE Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

MEDLINE 2238257

PUBMED 12477932

REFERENCE 2 (bases 1 to 2159)

AUTHORS Strausberg, R.

TITLE Direct Submission

JOURNAL Submitted (02-MAY-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>

COMMENT Contact: MGC help desk
Email: cgabp-r@mail.nih.gov
Tissue Procurement: DCTD/DTF
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc.mgc@nih.gov
Akhter, N., Ayale, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granice, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Iaric, P., Legaapi, R., Maduro, Q.L., Masello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Stanki, P., Thomas, P.J., Touchman, J.W., Tsugeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 47 Row: d Column: 2
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 24432061.
Location/Qualifiers
1..2159
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:59968 IMAGE:5467936"
/tissue types="Skin, melanoma, amelanotic"
/clone_lib="NIH_MGC_41"

FEATURES
source

1..2159
/lab_host="DH10B-R"
/note="Vector: pOTB7"
1..2159
/gene="FLJ32203"
/db_xref="LocusID:130557"
201..1826
/codon_start=1
/product="FLJ32203 protein"
/protein_id="AAH52282.1"
/db_xref="GI:30354480"
/db_xref="LocusID:130557"
/translation="MPRRKQSHPOVKCEGVKVDTEDSLDEGPALVLESDLLGLQDL
EPEEESEEGDNGSLGFRDSEGLSGARPGLPGLSDDESGGSGRAESAESVE
EPAGPEARERGRPACOLCGPTGEGCCGAGGGGPGPLPPLLYSCLCTFFVSH
YSHLKRHMTGSGERKPFRCGRCPYASAGLVNLRIRHTGKPKYRCHPCFPACSSL
GNLRHORTHAGPTTPTCTGFCCTPRPAPSPPTGEGAVPRPPEADALLPDLUL
HVPFGGSLPDCQURGEGLCGTSEPLLELFPWTGCGGCGLEGEGRGAA
MCRCMRGAGGAGGCGPQSDKPGFACSLCPFATHPNLHARHMKHSGEKPRCAR
CPYASHLNLRKHQVHTGKPKYKPCLPYACGNLANLRKHGRHSGDKPRFCSICN
YSCNOSMLKHEMLRHTGKPKFCATCAVTTGHDVYKHKQVHGGAGGPGLSASE
GNAPRSPSPVLSRRPFPALGTAGSRVHTDSS"
1440..1661
/note="COG5048; Region: COG5048, FOG: Zn-finger [General
function prediction only]"
/db_xref="CDD:COG5048"

misc_feature
1440..1661
/note="COG5048; Region: COG5048, FOG: Zn-finger [General
function prediction only]"
/db_xref="CDD:COG5048"

ORIGIN

Query Match 39.1%; Score 1505.8; DB 9; Length 2159;
Best Local Similarity 88.5%; Pred. No. 0;
Matches 1716; Conservative 0; Mismatches 178; Indels 44; Gaps 6;

QY 1917 AGTGATCTGAAGATTCCTTCGACGAAGTCTCTGGGGCCCTCGTGTGGAGAGCGGATTT 1976
DB 254 AGTGATCTGAAGATTCCTTCGACGAAGTCTCTGGGGCCCTCGTGTGGAGAGTATTT 313

QY 1977 GCTACTAGGCGAAGATCTGGAGTTTGAAGAGAGAGAGAGAGAGTGAAGTGCAGGCCA 2036
DB 314 GCTACTAGGCGAAGATCTGGAGTTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 373

QY 2037 CAACGACAGCTCATGGGCTTTGAGAGAGACTCTGAAGAGAGACTCTCAGGGGGGCGAGACC 2096
DB 374 CAGTGACAGCTCATGGGCTTTGAGAGAGACTCTGAAGAGAGACTCTCTGGGGGCGAGGCC 433

QY 2097 TGGACTTCCCTATGGGCTGAGTACGACGAGTCTGGGGGGCGGCGCCACCTAAGTGCAGGA 2156
DB 434 TGGGCTTCCCTATGGGCTGAGTACGACGAGTCTGGGGGGCGGCGCCACCTAAGTGCAGGA 493

QY 2157 CAGTGAGTTGAGGAGCCAGCCAGGGTCCAGGGGGGCGGCGGAGGAGAGAGAGAGAGAGAG 2216
DB 494 GAGTGAGTTGAGGAGCCAGCCAGGGTCCAGGGGGGCGGCGGAGGAGAGAGAGAGAGAG 553

QY 2217 AGCCTGTGAGTGTGTGGGGGGCGGCGAGAGTGTGGGGGGCGGCGGAGAGAGAGAGAGAG 2276
DB 554 AGCCTGTGAGTGTGTGGGGGGCGGCGAGAGTGTGGGGGGCGGCGGAGAGAGAGAGAGAG 613

QY 2277 GGGTGG 2336
DB 614 GGGTGG 673

QY 2337 GTCCCACTACTCTGAGAGCCACCTGAGAGCGGCGCATGACAGACACAGAGCGGGGAGAGCGGTT 2396
DB 674 GTCCCACTACTCTGAGAGCCACCTGAGAGCGGCGCATGACAGACACAGAGCGGAGAGCGGTT 733

QY 2397 CGGTGTGG 2456
DB 734 CGGTGTGG 793

QY 2457 CACCCATCTGCGGAG 2516
DB 794 CACCCATCTGCGGAG 853

QY 2517 GGGCAACTGAGGGGGGAG 2576

854	GGGCAACTGAGCGGCATCAGGTACCAACGAGGGCCCCCACTCTCCCTGCCCGAC	913
2577	CTGTGGCTTTTCGATGTGTCTCCACGACAAACCGGCGCTCCAGTCCACAGACAGA	2636
914	CTGTGGCTTCGGCTGTGTACTCCACGACGAGCCGGCTCCAGTCCACAGACAGA	973
2637	GGGGCAAATGCCCCGACGATCAGAAATGGCTGTATCTTCGCCAGACTTGATCTTCATCT	2696
974	GGGGGGGGTGCCTCCCGCAGCCTGAAGATGCTCTGTCTTCCTTCAGATTGAGCCTCCATGT	1033
2697	GCCACAGGTGGTGCACAGTTTCTGCGCAGACTGTGGCGAGCTCGGGGTGAAGGGGAGAG	2756
1034	GCCACAGGTGGTGCACAGTTTCTGCGCAGACTGTGGCGAGCTCGGGGTGAAGGGGAGGG	1093
2757	CTTGTGTGAATCTGGATTCGAAACCACTGCACAGACTACTGTTCCTTGGACCTGCCGGG	2816
1094	CCTCTCGGAGCTGGATCAGAAACCACTGCCAGAGCTGCTATTCCCTTGGACCTGCCGGGG	1153
2817	CTGTGACAGAACTGGAGAGGGTGAAGGCACAGCGCTGGGAGCTGCATGTGTGGGGC	2876
1154	CTGTGACAGAGCTGGAGAGGGTGAAGGTAGTGGCTGGGAGCTGCCATGTGTGGGGC	1213
2877	CTGCATGCGAGGAGAGCTGGAGGGTGTGCACTGGGGGACCCCAAGGCCCTTGGTGACAA	2936
1214	CTGCATGCGAGGAGAGCTGGAGGGGTGCAGTGGGGGGCCCCAGGGCCCCAGTGACAA	1273
2937	AGCCTTGTGCTGTAGTTTATGCCCTTTGCGCACTCACTACCCCAACCACTGGCTCGGCA	2996
1274	AGCCTTGTGCTGTAGCCTCTGCCCTTTGGCATCTCACTATCCCAACCACTGGCCCGGCA	1333
2997	CATTGAAGACTCACAGTGGTGAGAAACCTTTCCGCTGTGCCCGTGTCCATPAGCCTCTGC	3056
1334	CATTGAAGACACAGTGGTGAGAAAGCCCTTCGCTGGCGCCGTGTCTTATGCTCTCTGC	1393
3057	TCACTGTGATAACTGTAAACGGCACACAGCGGCTGCCACAGAGGAAAGCCCTACAAGTG	3116
1394	TCACTGTGATAACTGTAAACGGCACACAGCGGCTGCCATACAGAGAGAGGCCCTACAAGTG	1453
3117	CCCCCTCTGCTATGCCCTGTGGCAACTGGCGCAACCTCAAGCGTCATGGTCCGATCCA	3176
1454	CCCCCTCTGCCCTTATGCTGTGGCAATCTGGCCACCTCAAGGTCATGGTCCGATCCA	1513
3177	CTCTGTGACAAACCTTTTCGTTGTAGCTTTGGAACTACAGCTGMAACAGAGTATGAA	3236
1514	CTCTGTGACAAACCTTTTCGTTGTAGCTTTGGCACTACAGCTGCAACACAGAGCATGAA	1573
3237	CCTCAACGTCATATGCTGCGACACACGCGGAGAGGCCCTTCGGTGTGCCACCTGCGC	3296
1574	CCTCAACGTCATATGCTGCGGACACAGCGGAGAGGCCCTTCGGTGTGCCACCTGCGC	1633
3297	CTATACCAAGGCCACTGGGACAACTACAGCGTCATCAGAAGGTGCATGGCCATGGTGG	3356
1634	CTATACCAAGGCCACTGGGACAACTACAGCGCCACAGAGGCCATGGCCACCGTGG	1693
3357	AGCAGGAGCCCTGGTCTCTCTGCCCTGAGGGCTGGGCCCACTCATAGCCCAACCTC	3416
1694	GGCAGGAGCCCTGGTCTCTCTGCCCTGAGGGCTGGGCCCACTCATAGCCCAACCTC	1753
3417	TGTTTTGAGCACTCGGGGTCCAGCAGCCCTGGGTGTACTGTGTAGCAGGCTCTTCAATC	3476
1754	TGTTTTGAGCTCTCGGGGCCCAACAGCCCTGGGAGCTGCTGGCAGCGGCTGTGCACAC	1813
3477	AGACTCACTTGAACTAACTAGTTCTTTTACCTGGGGCTCTAGGAATTAGCCCTATGCC	3536
1814	AGACTCACTTGAACTAACTAGTTCTTTTACCTGGGGCTCTAGGAATTAGCCCTATGCC	3596
3537	TCCTGCAATTTTATACAAATGAACTAGAAAAACACTTTTCCCTTTCTCCCGCGCTGGTCAGG	3596
1841	CCCCATGTTTTATACAGACGACCAAGAGCCACTTT-----TTTCTCCCGCTGGCCAGG	1896
3597	GGCTCCAACAGACTAACTTAGGCATCTATATGACACAGCTGTAATCCATGTGTACGGGG	3656
1897	GGCTCCAACAGACTAACTTAGGCATCTATATGACACAGCTGTAATCCATGTGTGGGGGGG	1956

Qy	3657	-CATATAGACCAAGGGA-CTTGCTTTAGCTCAGGTACCAGATGAGTAACTGATTAGGG	371.
Db	1957	CCCATATGGACCAAGGGAACCTTCCTTTGACTGAGGCATCTCATGAGCTCAGTGAGAGGG	2016
Qy	3715	CCVTGGATTACCGGCCACTGCTCCCAGAGGCTATGGATGAACCTGGTTGGGAG-CTGCCCCA	3773
Db	2017	CCCTGTATTCACTCCACTGCCCCAGGGGCTGTGGACAAACCGCTGGGGGACTGCCCA	2076
Qy	3774	GCCTTTTTAC-----TGTTTTTAACCTTATTTCAGTGCCTTTATAATAAGAAACA	3829
Db	2077	GCCTCCCACTGTTTATTAACTTATTTCAGTGCCTTTATAATAAGAAACA	2136
Qy	3830	AAAAAAAAAAAAAAA 3847	
Db	2137	GCCAAAAA 2154	
RESULT 11			
AX274856			
LOCUS			
DEFINITION Sequence 121 from Patent WO0172777.			
ACCESSION AX274856			
VERSION AX274856.1 GI:16547488			
KEYWORDS			
SOURCE Homo sapiens (human)			
ORGANISM Homo sapiens			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE 1 Hillman,J.L., Baughn,M.R., Yue,H., Lal,P., Lu,D.A., Patterson,C.,			
Azimzai,Y., Bandman,O., Tang,Y.T., Mathur,P., Shah,P., Au-Young,J,			
and Reddy,R.			
TITLE Transcription factors			
JOURNAL Patent: WO 017277-A 121 04-OCT-2001;			
Incyte Genomics, Inc. (US)			
FEATURES			
source			
1..1975			
Location/Qualifiers			
/organism="Homo sapiens"			
/mol_type="unassigned DNA"			
/db_xref="taxon:9606"			
/note="Incyte ID No: 870100CB1"			
ORIGIN			
Query Match 38.9%; Score 1498; DB 6; Length 1975;			
Best Local Similarity 88.6%; Pred. No. 0;			
Matches 1707; Conservative 0; Mismatches 176; Indels 44; Gaps 6;			
Qy	1917	AGTGGATCTGAAGATTCCTTCGACGAAGTGCTCTGGGGCCCTGGTGTGGAGACCGATTT	1976
Db	86	AGTGGATCTGAAGATTCCTTCGACGAAGGACCCGGGGCCCTGGTATGGAGAGTGATTT	145
Qy	1977	GCTACTAGGCCAACATCTGGAGTTTGAAGAGGAAGAGAAGATCAAAGTACGGCCA	2036
Db	146	GCTACTAGGCCAACATCTGGAGTTTGAAGAGGAAGAGAAGAGGAGAAAGCGACGCA	205
Qy	2037	CACAGCACAGCTCATAGGGCTTTGAGAGAGACTCTGAAGAGAGACTCTCAGGGGGCCAGACC	2096
Db	206	CAGTGACCACTCATAGGGCTTCGAGAGAGACTTCGAAAGAGAGACTCTCTGGGGGGCCAGGCC	265
Qy	2097	TGGACTTCCCTATGGGCTGAGTAGCAGCAGTCTGGGGGGCGCGCACTAAAGTGGGA	2156
Db	266	TGGGCTTCCCTATGGGCTGAGTAGCAGCAGTCTGGGGGGCGCGCACTAAAGTGGGA	325
Qy	2157	GAGTGAAGTTGAGGAACACGCCAGGGGTCCAGGGGAGGCCAGGGGTGAGAGGCCAGGCC	2216
Db	326	GAGTGAAGTTGAGGAGCCAGCCAGGGGTCCAGGGGAGGCCAGGGGTGAGAGGCCAGGCC	385
Qy	2217	AGCCTGTCACTGTGTGGGGGGCCAGAGGTGAGGGGCCGTGTTGTGGGGCAGGAGGGCG	2276
Db	386	AGCCTGCCAGCTGTGTGGGGGGCCAGAGGTGAGGGGCCGTGTTGTGGGGCAGGAGGGCC	445
Qy	2277	GGGTGGGGGGCCCCCGCTGCCCCACCGTTACTACTATCAATGCCGGCTGTGCGCTTCGT	2336

1526 GGCAGAGGGCTGTGCTCTCTGCGCTCTGAGGGCTGGGCCCCACCTCATAGCCCCCCTC 1586
3417 TGTTTTGACACTCGGGGTCCAGCAGCCCTGGGTGTATCTGGTAGCAGGGCTCTTCATTC 3476
1586 TGTTTTGAGCTCTCGGGGCCACACGAGCCCTGGGACTGTCTGGCAGACGGGCTGTCCACAC 1645
3477 AGACTCACCTTGAACATACTAGGTTCTTTTACCTGGGGCTCTAGGAATTAGCCCTATGCC 3536
1646 AGACTCATCTGAAC-----TAGTCTCTTCTT 1672
3537 TCCTGCATTTTATACAAATGAAGTACAGAAACCACTTTCCCTTTCTCCCGCGCTGGTCAGG 3596
1673 CCCCATGTTTATACAGCGGACCAAGCCACCTT---TTTCTCCCGCGCTGGCCAGG 1728
3597 GGCTCCACACAGACTAACTAGGCACATATAGCAGCAGCTGTAATCCATGGTCAAGGGGG 3656
1729 GGCTCCACACAGACTAACTAGGCACATATAGGACCAAGCAGCCCACTGGGCGGGGGGG 1788
3657 -CCATATAGACCAGGGGA-CTTGCTCTTAGCTCACGTACAGATGAGTAAGTCAATTAGGG 3714
1789 CCATATGACACAGGGGACCTTGCTTGACTGAGGCACCTCACGAGCTCAGTGAGAAGG 1848
3715 CTTTGATTCACCGCACCTGCTCCAGAGGCTATGATGAACTGGTTGGAG-CTGCCCA 3773
1849 CCCTGTATTCACCTCCACTGCTCCCGCAGGGGCTCTGGACAAACCGGCTGGGGGACTGCCCA 1908
3774 GCCTTTTAC---TGTTTAACTATTTCAGTCTCTTATAATAAGAAACACTAACAGA 3829
1909 GCCTCCACCTGTTATTTAACCTATTTCAGTCTCTTATAATAAGAAACACTAACAAA 1968
3830 AAAAAA 3836
1969 AAAAAA 1975

RESULT 12
BD242880
LOCUS 1996 bp DNA linear PAT 17-JUL-2003
DEFINITION Secreted proteins and polynucleotides encoding them.
ACCESSION BD242880
VERSION BD242880.1 GI:33052650
KEYWORDS JP 2002536973-A/31.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutharia; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 1996)
Valenzuela,D., Yuan,O., Hoffman,H., Hall,J. and Rapiejko,P.
Secreted proteins and polynucleotides encoding them
Patent: JP 2002536973-A 31 05-NOV-2002;
ALPHABASE INC
OS Homo sapiens (human)
PN JP 2002536973-A/31
PD 05-NOV-2002
PF 18-FEB-2000 JP 2000599860
PR 19-FEB-1999 US 60/120680,23-APR-1999 US 09/298733 PR
17-AUG-1999 US 60/149639,23-SEP-1999 US 60/155686 PR
01-OCT-1999 US 60/157247,29-NOV-1999 US 60/167823 PR
29-NOV-1999 US 60/167822,15-FEB-2000 US 60/182711 PI DARIO
VALENZUELA, OLIVE YUAN, HEIDI HOFFMAN, JEFF HALL, PETER PI, RAPIEJKO
PC C12N15/09,A61K38/00,A61P3/10,A61P5/14,A61P11/00,A61P11/06,PC
A61P19/02,
PC A61E21/04,A61P25/14,A61P27/02,A61P29/00,A61P31/04,A61P31/10,
PC A61P31/12,
PC A61P31/18,A61P31/20,A61P31/22,A61P37/00,A61P37/06,C07K14/435,
PC C12N5/10,
PC C12P19/34//C12P19/34,C12R1:91,C12N15/00,C12N5/00,A61K37/02
CC Secreted proteins and polynucleotides encoding them FH Key
CC Location/Qualifiers
FT source
1..1996
/organism='Homo sapiens (human)'.
FT Location/Qualifiers
1..1996
source

Query Match	35.5%	Score 1367.6	DB 6	Length 1996	
Best Local Similarity	86.2%	Pred. No. 0			
Matches 1598	Conservative	0	Mismatches 210	Indels 45	Gaps 6
QY	2003	AAGAGGAAGAGGAGGATGAAGAGGTGACGGGCCACACGACACGACAGCTCATGGGCTTTGAGA	2062		
DB	139	AAGTGAAGAAGGGAGAGTTTGGAGCCTGATCCTACCCATGCTGATGCTCTCTCTTATGT	198		
QY	2063	GAGACTCTGAAGAGACTCTCAGGGGCCAGACCTGSACTTCCCTATGGGCTGAETGACG	2122		
DB	199	CTATTTTACACAGAGAGACTCTCTGGGGGCCAGGCGCTGGGCTTCCCTATGGGCTCAGCGACG	258		
QY	2123	ACGAGTCTGGGGGGCGGCGGCCACTAAGTGCAGAGAGTGAAGTTTGAGGAACACGACGAGGG	2182		
DB	259	ATGAGTCTGGGGCGGCGGCCACTAAGTGCAGAGAGTGAAGTTTGAGGACCGACGACGAGG	318		
QY	2183	GTCCAGGGAGGCCAGGGGTGAGAGGCCAGGCCAGCCTGCTGAGCTGTGTGGGGGCCGA	2242		
DB	319	GTCCAGGGAGGCCAGGGGTGAGAGGCCAGGCCAGCCTGCGAGCTGTGTGGGGGGCCGA	378		
QY	2243	CAGGTGAGGGGGCGGTGTGTGTGGGCGAGGAGGGGGTGGGGGGCCCCCGCTGCCCCCAC	2302		
DB	379	CAGGTGAGGGGGCGGTGTGTGTGGGCGAGGAGGGCGGGTGGGGGGCCCCCTGTGCCCCCAC	438		
QY	2303	GGTTACTGTACTCATGCCGCTGTGCCCTTTCGTGTCCTCCACTACTCAGAGCACCTGAGC	2362		
DB	439	GGCTACTGTACTCATGCCGCTCTGCACCTTCGTGTCCCACTACTCAGAGCACCTGGAAC	498		
QY	2363	GGCACATGCGACACACAGCGGGAGAACCGTTCCGCTGTGGCCGTGCCCATACGCT	2422		
DB	499	GGCACATGCGACACACAGCGGGAGAACCGTTCCGCTGTGGCCGTGCCCATACGCT	558		
QY	2423	CAGCCAGTTTCGTCAACTCAGCGGACATACCCGACCCACTACTGCGGAGAGCCCTACC	2482		
DB	559	CAGCCAGTCTGTCAACTGACAGCATATCCCGACCCACTGCGGAGAGCCCTACC	618		
QY	2483	GTGTGCCCCACTGCCCTTTGCCCTGCGAGCGCTGGGGCAACTGAGCGGGCATCAGCGCA	2542		
DB	619	GCTGTCCCCACTGTGCCCTTTGCCCTGCGAGCGCTGGGGCAACTGAGCGGGCATCAGCGTA	678		
QY	2543	CCACACAGGGGCTCCCACTCTCCCTGCCCAACCTGTGGCTTTTCGATGCTGTCTCCAC	2602		
DB	679	CCACGCGAGGGCCCCCACTCTCTCCCTGCCGACCTGTGGCTTCGCTGTGTACTCCAC	738		
QY	2603	GACCAACCCGGGCTCCGAGTCCCAAGAGGAGGGGCAATGCCCCGACGATCAGAA	2662		
DB	739	GACCAAGCCGGGCTCCAGTCCCAAGAGGAGGGGGGGGGTCCGCCGGACCTGAAG	798		
QY	2663	ATGGGCTGATCTCTGCCAGACTTGAGTCTTCATGTGCCACAGGTGTGCCAGTTTCTCTGC	2722		
DB	799	ATGCTCTGCTCTCTCCAGATTGAGCCTCCATGTGCCACAGGTGTGTGCCAGTTTCTCTGC	858		
QY	2723	CAGACTGTGGGACGCTCGGGGGTGAAGGGGAGAGCTTGTGTGAACTGTGATTCGAAACAC	2782		
DB	859	CAGACTGTGGGACGCTCGGGGGTGAAGGGGAGGGCCCTCTCGGGACCTGGATCAGAAACAC	918		
QY	2783	TGCCAGAGCTACTGTCTCCCTTGGACCTGCCGGGCTGTGGACAGAACTGGAGAGGGTG	2842		
DB	919	TGCCAGAGCTGCTATTTCCTTGGACCTGCCGGGCTGTGGACAGAGCTGGAGAGGGTG	978		
QY	2843	AGGCGAGAGGCTGGAGCTGCCATGTGTGGGCGCTGCATCCGAGGAGAGGCTGGAGGGG	2902		
DB	979	AGGCTAGTCGGCTGGAGCTGCCATGTGTGGGCGCTGCATCCGAGGAGAGGCTGGAGGGG	1038		
QY	2903	TTGCCACTGGGGGACCCCGAGGGCCCTGTGTGACAAAGGCTTTTGCCTGTAGTTTATGCCCT	2962		
DB	1039	GTGCCAGTGGGGGGCCCCAGGGCCCCAGTGACAAAGGCTTTTGCCTGTAGCTCTGCCCT	1098		

AC074117 195880 bp DNA linear PRI 21-FEB-2000
Homo sapiens BAC clone RP11-413M20 from 2, complete sequence.
AC074117
AC074117.7 GI:18464319
HTG.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 195880)

AUTHORS Sulston, J.E. and Waterston, R.
 TITLE Toward a complete human genome sequence
 JOURNAL Genome Res. 8 (11), 1097-1108 (1998)
 MEDLINE 99063792
 PUBMED 9847074
 REFERENCE 2 (bases 1 to 195880)
 AUTHORS Sun, H., Kozlowicz, A. and Dignan, G.
 TITLE The sequence of Homo sapiens BAC clone RP11-413M20
 JOURNAL Unpublished (2001)
 REFERENCE 3 (bases 1 to 195880)
 AUTHORS Waterston, R.H.
 TITLE Direct Submission
 JOURNAL Submitted (14-JUL-2000) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
 REFERENCE 4 (bases 1 to 195880)
 AUTHORS Waterston, R.H.
 TITLE Direct Submission
 JOURNAL Submitted (05-FEB-2002) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
 REFERENCE 5 (bases 1 to 195880)
 AUTHORS Waterston, R.H.
 TITLE Direct Submission
 JOURNAL Submitted (08-FEB-2002) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 REFERENCE 6 (bases 1 to 195880)
 AUTHORS Waterston, R.H.
 TITLE Direct Submission
 JOURNAL Submitted (21-FEB-2002) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 REFERENCE 7 (bases 1 to 195880)
 AUTHORS Waterston, R.H.
 TITLE Direct Submission
 JOURNAL Submitted (08-FEB-2002) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 COMMENT On Feb 1, 2002 this sequence version replaced gi:14495402.
 ----- Genome Center
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: <http://genome.wustl.edu/gsc>
 Contact: sapiens@watson.wustl.edu
 ----- Summary Statistics
 ----- Center project name: H_NH0413M20

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

The RPC1-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Woon, P.Y., Zhao, B., Frengen, E., Tatenno, M., Cacanese, J.J. and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong

and coworkers at <http://www.chori.org>
 VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:

Actual start of this clone is at base position 1 of RP11-413M20; actual end is at base position 195880 of RP11-413M20.

A transposon exists in the growth of this clone which has been deleted from the finished sequence. This transposon would insert after base 78210. Unsure bases exist between bases 165498 to 165505. Unresolved tandem repeat exists between bases 4700 to 6435. Polymorphisms have been identified between AC009427 and AC074117. Data from AC009427 was used to finish this clone, AC074117.

FEATURES

source	Location/Qualifiers
1..195880	/organism="Homo sapiens"
	/mol_type="Genomic DNA"
	/db_xref="taxon:9606"
	/chromosome="2"
	/map="2"
	/clone="RP11-413M20"
	/clone_lib="RPC1-11"
1..207	/rpt_family="L1"
208..500	/rpt_family="Alu"
501..641	/rpt_family="L1"
746..1042	/rpt_family="Alu"
1221..1538	/rpt_family="Alu"
1518..1538	/rpt_family="Alu"
1546..1866	/rpt_family="(A)n"
2505..3520	/note="similar to Homo sapiens EST AW327943 (NID:G6798438)"
2507..2520	/note="similar to Homo sapiens EST AW327944 (NID:G6798439)"
dr03a10.y1	/note="similar to Homo sapiens EST AW327944 (NID:G6798439)"
2722..2766	/rpt_family="AT-rich"
2959..3251	/rpt_family="Alu"
3153..3202	/note="similar to Homo sapiens EST AU130518 (NID:G10990872)"
3558..3645	/rpt_family="MER1_type"
3646..3952	/rpt_family="Alu"
3651..3671	/rpt_family="(TTTTTA)n"
3953..3995	/rpt_family="MER1_type"
3962..3990	/rpt_family="AT-rich"
3996..4284	/rpt_family="Alu"
4285..4331	/rpt_family="MER1_type"
4515..4593	/rpt_family="L2"
4826..4670	/rpt_family="(TCTCCC)n"
4780..6423	/note="CPG island (GC=74.2, o/e=0.70, #CpGs=131)"
5543..5910	/rpt_family="SVA"
5926..5950	

```

/rpt_family="GC_rich"
repeat_region 5988..6960
/rpt_family="SVA"
repeat_region 6293..6345
/rpt_family="GC_rich"
repeat_region 6942..6971
/rpt_family="A)n"
repeat_region 7085..7209
/rpt_family="L2"
repeat_region 7210..7518
/rpt_family="Alu"
repeat_region 7646..7800
/rpt_family="MER1_type"
repeat_region 9057..9346
/rpt_family="Alu"
repeat_region 9899..10049
/rpt_family="L2"
repeat_region 10029..10071
/rpt_family="(TA)n"
repeat_region 10107..10354
/rpt_family="L1"
repeat_region 10355..10533
/rpt_family="MER1_type"
repeat_region 10540..10886
/rpt_family="L1"
repeat_region 10887..11178
/rpt_family="Alu"
repeat_region 11179..11483
/rpt_family="L1"
repeat_region 11461..11491
/rpt_family="AT_rich"
repeat_region 11484..11602
/rpt_family="Alu"
repeat_region 11603..12618
/rpt_family="L1"
repeat_region 12619..12692
/rpt_family="ERV1"
repeat_region 12673..13130
/rpt_family="ERV1"
repeat_region 13145..13221
/rpt_family="L1"
repeat_region 13237..13292
/rpt_family="L1"
repeat_region 13293..13408

Query Match 32.6%; Score 1255.2; DB 9; Length 195880;
Best Local Similarity 82.2%; Pred. No. 3,7e-291;
Matches 1586; Conservative 0; Mismatches 204; Indels 139; Gaps 7;

QY 2003 AAGAGGAGGAGGAGGATGAAGGTGACGGCCACAGCCAGCTCATGGGCTTTGAGA 2062
DB 166942 AAGTGAAGGGAGAGATTGGAGCCTGATCCTACCTATGCTGATGCTCTCTTTATGT 167001
QY 2063 GAGACTCTGAAGGAGACTCTCAGGGGCCAGACTGGAATTCCTATGGCTGAGTGACG 2122
DB 167002 CTATTTCAACAGGAGACTCTCTGGGGCCAGGGCTGGCTTCCCTATGGCTGAGCGAG 167061
QY 2123 ACAGTCTGGGGCGCGCGCACTAAGTGGGAGTGAAGTTGAGGAACCAAGCCAGGG 2182
DB 167062 ATGAGTCTGGGGCGCGCGGCACTAAGTGGGAGTGAAGTTGAGGAGCCAGCCAGGG 167121
QY 2183 GTCCAGGGGAGGCGCAGGGGTGAGAGGCCAGGGCCAGCTGTGAGCTGTGTGGGGGCCGA 2242
DB 167122 GTCCAGGGGAGGCGCAGGGGTGAGAGGCCAGGGCCAGCTGTGAGCTGTGTGGGGGCCGA 167181
QY 2243 CAGGTGAGGGCCGTGTGTGGGGCAGAGGGCGGGGTGGGGGGCCCCCGCTGCCCGCCAC 2302
DB 167182 CAGGTGAGGGCCGTGTGTGTGGGGCAGAGGGCGGGGTGGGGGGCCCCCGCTGCCCGCCAC 167241
QY 2303 GGTACTGTACTCATGCGGGGTGTGCGCTTTTCGCTGTCCTACTACTCGAGCCACTGAAGC 2362
DB 167242 GGCTACTGTACTCATGCGGGCTGTGCACTTGTGCACTTGTGCACTTGTGCACTTGTGCACT 167301

```

```

QY 2363 GGCAATGTCAGACACACAGCGGGAGAGCCGTTTCGCTGTGGCCGCTGCCCATACGCT 2422
DB 167302 GGCAATGTCAGACACACAGCGGGAGAGCCGTTTCGCTGTGGCCGCTGCCCATACGCT 167361
QY 2423 CAGCCAGTTTGTCAACCTGACCGCATACCCGACCCCATATATGCGGAGAAAGCCCTTACC 2482
DB 167362 CAGCCAGTTTGTCAACCTGACCGCATACCCGACCCCATATATGCGGAGAAAGCCCTTACC 167421
QY 2483 GTTGTCCCACTGCCCTTTGCTGTGAGCAGCTGGGCAACCTGAGCGGCATCAGCGCA 2542
DB 167422 GCTGTCCCACTGCCCTTTGCTGTGAGCAGCTGGGCAACCTGAGCGGCATCAGCGTA 167481
QY 2543 CCCACAGCGGCTTCCCACTTCTCCCTGSCCAACCTGTGCTTTTCGATGTGTCTCCAC 2602
DB 167482 CCCACAGCGGCTTCCCACTTCTCCCTGSCCAACCTGTGCTTTTCGATGTGTCTCCAC 167541
QY 2603 GACCAACCGGCTTCCCACTTCCCACTGAGGAGGAGGAGCAANTGCCCGCAGTACAGA-- 2660
DB 167542 GACCAACCGGCTTCCCACTTCCCACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 167601
QY 2661 ----- 2660
DB 167602 GTAAGACACACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 167661
QY 2661 -----AATGCGCTGATCTGCGCAGCTTGAG 2687
DB 167662 GATTGGATTATAGCCCAAGGCTTTGTGCCACAGATGCTCTCTCCATGATTTGAG 167721
QY 2688 TCTTCATGTGCCACAGGTGGTGCAGTTTCTGCCAGAGCTGTGGGCGAGCTGCGGGGTGA 2747
DB 167722 CTTCCATGTGCCACAGGTGGTGCAGTTTCTGCCAGAGCTGTGGGCGAGCTGCGGGGTGA 167781
QY 2748 AGGGAGAGCTTGTGTGAACCTGATCCGAACCTGCGCAGAGCTACTGTTCCTTTGAC 2807
DB 167782 AGGGAGAGGCTTGTGTGAACCTGATCCGAACCTGCGCAGAGCTACTGTTCCTTTGAC 167841
QY 2808 CTGCGGGGCTGTGACAGGAACTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2867
DB 167842 CTGCGGGGCTGTGACAGGAACTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 167901
QY 2868 GTGTGGGCGCTGCATGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2927
DB 167902 GTGTGGGCGCTGCATGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 167961
QY 2928 TGGTGAACAGGCTTGTGCTGTGATGCTTATGCTTGTGCTTGTGCTTGTGCTTGTGCTTGT 2987
DB 167962 CAGTGAACAGGCTTGTGCTGTGATGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGT 168021
QY 2988 GGCTGGGCACTGAAGACTCACTGAGTGTGAGAAACCTTTCGCTGTGCTGCTGTGCTGCTTGA 3047
DB 168022 GGCTGGGCACTGAAGACTCACTGAGTGTGAGAAACCTTTCGCTGTGCTGCTGTGCTGCTTGA 168081
QY 3048 GCGCTTGTCTCATCTGGATTAACCTGAAACCGCAGCGGCTGCCACACAGAGAGAAAGCC 3107
DB 168082 TGCTTGTCTCATCTGGATTAACCTGAAACCGCAGCGGCTGCCACACAGAGAGAAAGCC 168141
QY 3108 CTACAAGTGTGCGGCTTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 3167
DB 168142 CTACAAGTGTGCGGCTTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 168201
QY 3168 TGGCTTGTCTCATCTGGTGAACAACTTTCGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 3227
DB 168202 TGGCTTGTCTCATCTGGTGAACAACTTTCGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 168261
QY 3228 GAGTATGAACCTCAAACTGATGCTGCGACACACAGCGGCGAGAGGAGGAGGAGGAGGAGGAG 3287
DB 168262 GAGTATGAACCTCAAACTGATGCTGCGACACACAGCGGCGAGAGGAGGAGGAGGAGGAGGAG 168321
QY 3288 CACCTGCGGCTTATACCAAGGCGCTTGGGAGCAACTGACAGGCTCATCAGAGAGGTCATGG 3347
DB 168322 CACCTGCGGCTTATACCAAGGCGCTTGGGAGCAACTGACAGGCTCATCAGAGAGGTCATGG 168381
QY 3348 CCATGTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3407

```


Db 721 GAAACGGCACCAGCGCTCCATACAGGAGAGCCCTACAAAGTGCCTCCCTCTGCCCCCTTA 780
 Qy 3132 TGCTCTGGCAACCTGGCCAACTCAAGCGTCAATGTGTCATCCACTCTGTGTGACAAACC 3191
 Db 781 TGCTCTGGCAATCTGGCCAACTCAAGCGTCAATGTGTCATCCACTCTGTGTGACAAACC 840
 Qy 3192 TTTTTCGGTGTAGCTTTTGGCACTTACAGCTGCAACAGAGATGTAACCTCAAAAGCTCATAT 3251
 Db 841 TTTTTCGGTGTAGCTTTTGGCACTTACAGCTGCAACAGAGATGTAACCTCAAAAGCTCATAT 900
 Qy 3252 GCTGCGACACACGGGCGAGAGCCCTTCGGCTGTGCGCACTGCGGCTTATACACAGGCGCA 3311
 Db 901 GCTGCGGCAACACAGGCGAGAGCCCTTCGGCTGTGCGCACTGCGGCTTATACACAGGCGCA 960
 Qy 3312 CTGGGCAACTACAGCGTCAATCAGAGCTGCAAGCTGTCAGAGAGAGAGGCGCTGG 3371
 Db 961 CTGGGCAACTACAGCGGCGAGAGGCTTCCGCTGTGCGCACTGCGGCTTATACACAGGCGCA 1020
 Qy 3372 TCTCTCTGCCCCCTGAGGGCTGGGCCCCCACCCTCATAGCCACCTCTGTGTTTGTAGCACTCG 3431
 Db 1021 TCTCTCTGCCCCCTGAGGGCTGGGCCCCCACCCTCATAGCCACCTCTGTGTTTGTAGCTCTCG 1080
 Qy 3432 GGGTCCAGAGCGCTGGGCTCTACTGTGTAGCAGGCTCTTCAATTCAGACTCACCTTGAAC 3491
 Db 1081 GGGCCCCACAGCGCTGGGCTCTACTGTGTAGCAGGCTCTTCAATTCAGACTCACCTTGAAC 1140
 Qy 3492 TAACTAGGTTCTTTTACCTGGGCTCTAGGAAATAGCCCTTATGCTCTCTGCTGCTGCTATAC 3551
 Db 1141 TAACTAGGTTCTTTTACCTGGGCTCTAGGAAATAGCCCTTATGCTCTCTGCTGCTGCTATAC 1167
 Qy 3552 AATGAACTAGAAACACCTTTCT 3611
 Db 1168 AGACGGACCAAGAGCCACCTT-----TTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1223
 Qy 3612 AACCTAGGCACTATAGGACAGGCTCAATCCCATGCTGAGGGGGCC-ATATAGACGAG 3670
 Db 1224 AACCTAGGCACTATAGGACAGGCTCAATCCCATGCTGAGGGGGGGCCATATAGACGAG 1283
 Qy 3671 GGA-CTTGTCTTACTGCTACCTGACAGATGAGTAAAGTATGAGGCTTGGATTTACCGC 3729
 Db 1284 GGACCTTGGCTTACTGAGGCACTTCTAGCTCAGTGAAGAGGGGCTGTATTCACCTC 1343
 Qy 3730 CACTGCTCCAGAGGCTATGATGAATCTGTTGGAG-CTGCCAGGCTTTTAC---TG 3784
 Db 1344 CACTGCCCCAGGGGCTGTGGAACACCGCTGGGGGACTGCCAGGCTCCACCTGTTT 1403
 Qy 3785 TTTTAACTATTTTCACTGCTTTTATATTAAGAGAAACACTAACAGAAAAA 3842
 Db 1404 ATTTAACTATTTTCACTGCTTTTATATTAAGAGAAACACTAACAGAAAAA 1461

RESULT 15
 AX714442 1456 bp DNA linear PAT 15-APR-2003
 LOCUS Sequence 1126 from Patent EP1293569.
 DEFINITION AX714442
 ACCESSION AX714442
 VERSION AX714442.1 GI:29889394
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1
 AUTHORS Isogai, T., Sugiyama, T., Otsuki, T., Wakamatsu, A., Sato, H., Ishii, S., Yamamoto, J. I., Isono, Y., Hio, Y., Otsuka, K., Nagai, K., Irie, R., Tamechika, I., Seki, N., Yoshikawa, T., Otsuka, M., Nagahari, K. and Masuho, Y.
 TITLE Full-length cDNAs
 JOURNAL Patent: EP 1293569-A 1126 19-MAR-2003;
 Helix Research Institute (JP); Research Association for
 Biotechnology (JP)
 FEATURES Location/Qualifiers
 source
 1..1456

ORIGIN
 Query Match 29.0%; Score 1115.2; DB 6; Length 1456;
 Best Local Similarity 87.4%; Pred. No. 1.3e-257;
 Matches 1305; Conservative 0; Mismatches 144; Indels 44; Gaps 6;
 Qy 2341 CACTACTCGAGCGCACTGAGCGGCATGACGACACACAGCGGGGAGAGCGTTCCTCCG 2400
 Db 1 CACTACTCGAGCGCACTGAGCGGCATGACGACACACAGCGGGGAGAGCGTTCCTCCG 60
 Qy 2401 TGTGGCGCTGCGCCATACGNTCAGCCCACTGTCTCAACCTGACGCGACATACCCGACC 2460
 Db 61 TGTGGCGCTGCGCCATACGCTCAGCCCACTGTCTCAACCTGACGCGACATACCCGACC 120
 Qy 2461 CATACTGGCGAGAGCGCTTACCGTGTCCCACTGCGCCCTTTGCTCTGACGAGCTGGGC 2520
 Db 121 CACACTGGCGAGAGCGCTTACCGTGTCCCACTGCGCCCTTTGCTCTGACGAGCTGGGC 180
 Qy 2521 AACCTGAGCGCGCATGAGCGCACCCACACAGGGGCTCCCACTCTCTCTCTGCCCCAACCCTGT 2580
 Db 181 AACCTGAGCGCGCATGAGCGTACCCACGAGGGGCCCCCACTCTCTCTCTGCCCCAACCCTGT 240
 Qy 2581 GCGTTTTCATGCTGTCTCAGACCAACCGCGCTCCAGTCCAGTCCACAGAGCAGAGGGG 2640
 Db 241 GCGTTTTCATGCTGTCTCAGACCAACCGCGCTCCAGTCCAGTCCACAGAGCAGAGGGG 300
 Qy 2641 ACAATGCCCGCCACCATCAGAAATGCGCTGATCTCTGCGAGACTTGAATCTTCAATGTGCCA 2700
 Db 301 GCGGTGCCCGCCCGACCTGAGATGCTCTGCTCTTCCAGATTTGAGCTTCCATGTGCCA 360
 Qy 2701 CAGGTGTGTGCGAGTTCTCTCCAGACTGTGGGAGCTGTGGGAGTGAAGGGGAGAGCTTG 2760
 Db 361 CCAGGTGTGTGCGAGTTCTCTCCAGACTGTGGGAGCTGTGGGAGTGAAGGGGAGAGGGCTC 420
 Qy 2761 TGTGAACTGTGATCGAACCACTGCGAGAGCTACTGTTCCCTTGGAGCTGCTGCGGGGCTGT 2820
 Db 421 TCGGGATGTGATCAGAACCACTGCGAGAGCTGCTATTCCTTGGAGCTGCTGCGGGGCTGT 480
 Qy 2821 GGACAGGAACCTGGAGAGGGTGGAGGAGCAGAGCTGGAGAGCTGCATGTGTGGGCGCTGC 2880
 Db 481 GGACAGAGCTGGAGAGGGTGGAGGAGTGGAGGAGCTGCATGTGTGGGCGCTGC 540
 Qy 2881 ATGGAGGAGGGCTGGAGGGTGGAGGAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2940
 Db 541 ATGGAGGAGGGCTGGAGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 600
 Qy 2941 TTTGCTGTAGTTTATGCCCCCTTTTGGCCACTCACTACCCCAACCCACTGGCTGGCGACATG 3000
 Db 601 TTTGCTGTAGCTCTGCCCCCTTTTGGCCACTCACTATCCCAACCCACTGGCGCGCACATG 660
 Qy 3001 AAGACTCAGTGTGAGAAACCTTCCGCTGTGCGCGCTGCTCCATAGCCCTCTGCTCAT 3060
 Db 661 AAGACACACAGTGTGAGAAACCTTCCGCTGTGCGCGCTGCTCTTATGCTCTGCTCAT 720
 Qy 3061 CTGGATAACCTGAAACCGCACCCAGCGCTCCACAGGAGAGAGAGAGAGAGAGAGAGAGAG 3120
 Db 721 CTGGATAACCTGAAACCGCACCCAGCGCTCCATACAGGAGAGAGAGAGAGAGAGAGAGAG 780
 Qy 3121 CTCTGTGCGTATGCTGTGGCAACCTGCGCAACTCAAGCGCTCATGTGCGATCCACTCT 3180
 Db 781 CTCTGCGCTTATGCTGTGGCAATCTGGCCCAACCTCAAGCGCTCATGTGCGATCCACTCT 840
 Qy 3181 GGTGACAAACCTTTTGGGTGTAGCCCTTTGCAACTACAGCTGCAACAGAGAGATGAACCTC 3240
 Db 841 GGTGACAAACCTTTTGGGTGTAGCCCTTTGCAACTACAGCTGCAACAGAGAGATGAACCTC 900
 Qy 3241 AAACGCTCATGCTCCGACACACAGGGCGAGAGAGCGCTTCCGCTGTGCGCACTGCGCCCTAT 3300
 Db 901 AAACGCTCATGCTCCGCGACACAGCGGAGAGAGCGCTTCCGCTGTGCGCACTGCGCCCTAT 960

/organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

Qy 3301 ACCACAGGCCACTCGGACAACTACAAGCGTCATCAGAAGGTGATGCCCATGTGTGAGCA 3360
 Db |||||
 Qy 961 ACCACGGGCCACTCGGACAACTACAAGCGGCCACACAGAAGGTGATGCCCATGTGTGAGCA 1020
 Db |||||
 Qy 3361 GAGGGGCTGTCTCTCTGCGCCCTGAGGGCTGGGCCCACTCATAGCCACCCCTCTGT 3420
 Db |||||
 Qy 1021 GAGGGGCTGTCTCTCTGCGCTCTGAGGGCTGGGCCCACTCATAGCCACCCCTCTGT 1080
 Db |||||
 Qy 3421 TTGAGCACTCGGGGTCCAGACGCCCTGGGTGCTACTGTAGCAGGGCTCTTCATTACAGAC 3480
 Db |||||
 Qy 1081 TTGAGCTCTCGGGGCCACACAGCCCTGGGACTGCTGGCAGCGGGCTGTCCACACAGAC 1140
 Db |||||
 Qy 3481 TCACCTTGAACCTAAGTCTAGGTCTTTTACCTGGGCTCTAGGAATTAGCCCTATGCCTCCT 3540
 Db |||||
 Qy 1141 TCATCCTGAACCTAGTCTCTTCTC-----CCC 1167
 Qy 3541 GCATTTTATACAAATGAACCTAGAACACCACTTTCCCTTTCTCCCGCTGTGTCAGGGGCT 3600
 Db |||||
 Qy 1168 ATGTTTATACAGACGGGACCAAGGCCACTT-----TTTCTCCCGCTGGCCAGGGGCT 1223
 Qy 3601 CCACACAGACTAACCTAGGCACTATATGGACCAAGCCTGAATCCCATGGTCAGGGGG-CCA 3659
 Db |||||
 Qy 1224 CCACACAGACTAACCTAGGCACTATATGGACCAAGCCTGAATCCCATGGTCAGGGGGGCCA 1283
 Qy 3660 TATAGACCAAGGGA-CTTGTCTTAGCTCACGTACAGATGAGCTAAGTATTAGGGGCTT 3718
 Db |||||
 Qy 1284 TATGACCAAGGGACCTTGGCTTGACTGAGGCACTTCAATGAGCTCAGTGAAGGGGCTT 1343
 Qy 3719 GGATTACCGGCACCTGTCTCCAGAGGCTATGGATGAACCTGGTTGGAG-CTGCCCCAGCCT 3777
 Db |||||
 Qy 1344 GTATTCACTCCACTGCCCCCAGGGGCTGTGGACAAACCGGCTGGGGGACTGCCCAGCCT 1403
 Qy 3778 TTTAC-----TGTTTAACTTATTTTCACTGCTTTTATATAAAGGAACACTAAC 3826
 Db |||||
 Qy 1404 CCCACCTGTTTATTTAACTTATTTTCACTGCTTTTATATAAAGGAACACTAAC 1456

Search completed: May 14, 2004, 03:12:30
 Job time : 14677 secs

D.T.B: 4/28/98

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 13, 2004, 15:37:55 ; Search time 251 Seconds
(without alignments)
8505.562 Million cell updates/sec

Title: US-09-673-994A-7
Perfect score: 3847
Sequence: 1 9gaggttagactttattc.....gaaaaaaaaaaaaaaaaa 3847

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:*
1: /cgn2_6/ptodata/2/ina/5A COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PCRUS COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	815.2	21.2	1558	4	US-09-620-312D-32
2	615.8	16.0	1824	2	US-08-822-701-1
3	615.8	16.0	1824	3	US-08-935-855-1
4	615.8	16.0	2081	3	US-08-935-855-21
5	270.2	7.0	1890	3	US-08-935-855-19
6	206.2	5.4	235	4	US-09-016-434-1049
7	113.8	3.0	2678	4	US-09-016-434-1229
8	111.8	2.9	2416	4	US-09-016-434-1058
9	110.4	2.9	524	4	US-08-833-381-427
10	100.4	2.6	3536	3	US-09-418-640-3
11	100.4	2.6	3720	1	US-08-074-967-1
12	100.4	2.6	3720	2	US-08-553-541B-1
13	100.4	2.6	3720	3	US-09-268-202-1
14	100.4	2.6	3720	5	PCT-US94-06669-1
15	96.4	2.5	2769	4	US-09-620-312D-309
16	94.8	2.5	2680	3	US-08-063-035-1
17	94.8	2.5	2992	4	US-09-362-123A-3
18	93	2.4	4468	4	US-09-620-312D-243
19	92	2.4	605	4	US-09-621-976-2809
20	92	2.4	1033	4	US-09-620-312D-500
21	90.2	2.3	639	4	US-09-833-381-1188
22	90	2.3	2666	4	US-09-881-578A-1
23	90	2.3	3186	4	US-09-016-434-1390
24	89.4	2.3	1892	2	US-08-933-750C-66
25	89.4	2.3	1892	3	US-09-234-613-66
26	87.6	2.3	644	4	US-09-833-381-819
27	85.8	2.2	128779	4	US-09-497-855A-38

c 27

ALIGNMENTS

RESULT 1
US-09-620-312D-32
; Sequence 32, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunging
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 32
; LENGTH: 1558
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (50)..(1174)
US-09-620-312D-32

Query Match 21.2%; Score 815.2; DB 4; Length 1558;
Best Local Similarity 85.7%; Pred. No. 1.4e-201;
Matches 1004; Conservative 0; Mismatches 123; Indels 45; Gaps 7;
QY 2674 CTGCCAGATTGAGTTTCATGTCGCCAGTGTGCCAGTTCCTGTCGACACTGTGGG 2733
DB 425 CCACAGATTGAGCTTCATGTCGCCAGTGTGCCAGTTCCTGTCGACACTGTGGG 484

Sequence 817, App
Sequence 1337, App
Sequence 689, App
Sequence 1399, App
Sequence 809, App
Sequence 163, App
Sequence 130, App
Sequence 586, App
Sequence 1406, App
Sequence 1401, App
Sequence 431, App
Sequence 172, App
Sequence 44, Appl
Sequence 1177, App
Sequence 911, App
Sequence 43, Appl
Sequence 821, App
Sequence 1054, App


```

RESULT 7
US-09-016-434-1229
; Sequence 1229, Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/016,434
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; PRIOR APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0002 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1229:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2678 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: g189043
US-09-016-434-1229

Query Match 3.0%; Score 113.8; DB 4; Length 2678;
Best Local Similarity 56.6%; Pred. No. 2.1e-19;
Matches 21; Conservative 0; Mismatches 162; Indels 0; Gaps 0;

Qy 2925 CCTGTGTGACAAAGCGTTCCTGTAGTTATTCGCCCTTTGGCACTCACTACCCACCA 2984
Db 1864 CACGGCGACAAAGTCCTTTGGCTGCGTGCAGTCCGGCGAGCGCTTGGCGCGCGCTCACT 1923
Qy 2985 CCTGGCTCGGCACATGAAGACTCACAGTGGTGAGAAACCTTCCGCTGTGCCCGCTGTCC 3044
Db 1924 GCTGTCTCAGCACCGGGCGGTGCACAGTGGCAGCGGCCCTTCCGCTGTGCCAGTGGCG 1983
Qy 3045 ATAGCGCTTGTCTATCTGTGATTAACCTGAAAACGGCACAGCGGCTCCACACAGGAGAAA 3104
Db 1984 CCAGAGCTTCGGCGACGGCTCCCAACCTGACGACACCGCGCGCATCCACCGGGGAGCG 2043
Qy 3105 GCCCTACAAGTGCCTCTGTCCGTATGCTGTGGCAACCTTGGCCAACTCAAGGCTCA 3164
Db 2044 GCCCTTCGCTTGCCGAGTGTGGCAGGCTTCCGCGAGGCGCTTACGCTCAGCGAGCA 2103
Qy 3165 TGGTGCATCCACTCTGGTGACAAAACCTTTTCGGGTAGCGTTTGGAACTACAGCTGCAA 3224
Db 2104 TCTCCGGGTACACACGGGCGAGAAAACCTTTTGGCTCGCCGAGTGTGGCAGCGCTTCAG 2163

```

QY 3225 CCAGAGTATGACCTCAACGTCATATGCTGGACACACGCGGAGAGCCCTTCCGCTG 3284
 Db 2164 CCAGGCGCTCAAGCTACCGCTCATCAGAGACACACCGCGGAGAGCCCTTCCGCTG 2223
 QY 3285 TGCCACCTGGCC 3297
 Db 2224 CGGTGAGTGGGC 2236

RESULT 8

US-09-016-434-1058
 ; Sequence 1058, Application US/09016434
 ; Patent No. 6500938
 ; GENERAL INFORMATION:
 ; APPLICANT: Janice Au-Young
 ; APPLICANT: Jeffrey J. Seilhamer
 ; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
 ; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
 ; NUMBER OF SEQUENCES: 1490
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
 ; STREET: 3174 PORTER DRIVE
 ; CITY: PALO ALTO
 ; STATE: CALIFORNIA
 ; COUNTRY: USA
 ; ZIP: 94304
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/016,434
 ; FILING DATE: HEREMITH
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE:
 ; CLASSIFICATION:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Zeller, Karen J.
 ; REGISTRATION NUMBER: 37,071
 ; REFERENCE/DOCKET NUMBER: PA-0002 US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (650) 855-0555
 ; TELEFAX: (650) 845-4166
 ; INFORMATION FOR SEQ ID NO: 1058:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 2416 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; IMMEDIATE SOURCE:
 ; LIBRARY: GENBANK
 ; CLONE: g1050529
 ; US-09-016-434-1058

Query Match 2.9%; Score 111.8; DB 4; Length 2416;
 Best Local Similarity 54.2%; Pred. No. 6.6e-19;
 Matches 227; Conservative 0; Mismatches 192; Indels 0; Gaps 0;
 QY 2930 GTGACAAAGCTTGGCTGTAGTTTATGCGCTTTGCGCACTCATACCCCAACCCCTGG 2989
 Db 761 GCGAGAAGGCTTACAAGTGGGATGATGCGCAAGGCTTCACTGGAGCACCACCTTC 820
 QY 2990 CTCGGACATGAAGACTCAGAGTGGTGAAGAACCTTTCGCTGTGCGCGTGTCCATAG 3049
 Db 821 TGAGACACCGGCGCATCCACCGGCGAGAGCCCTTCTCTCGGCGAGTGGGGAAG 880
 QY 3050 CCTCTGCTCATCTGGATAACCTGAAGCGGACCGCGCGTCCACACAGGAGAAAGCCCT 3109
 Db 881 CCTTCAGCTGCCACTGCTGCTCCTCAACGTGACAGGCGCATCCACACCGGCGAGCGCCCT 940

QY 3110 ACAAGTSCCCCTCTGTCCGTATGCTGTGGCAACCTGCGCAACCTCAAGCGTCAATGTC 3169
 Db 941 ACAAGTSCAGCGCTCGGAGAGGCGCTTCAGCTGCGAGCTGCTGTCTCAGCATGCACTGC 1000
 QY 3170 GCATCCACTCTGGTGAACAACTTTTCGGGTAGCTTTGGCACTACAGCTGCACCCAGA 3229
 Db 1001 GGGTGACACACCGGCGAGAGCCCTACCGGTGGGCGAGTSCGCGCAAGGCTTCAACCCAGC 1060
 QY 3230 GTATGAACCTCAAACTCATATGCTGCGACACACGCGGCGAGAGCCCTTCCGCTGTGCCA 3289
 Db 1061 GTACACCTTCAACCGCCACCGCATCCACGCGGCGAGAGCCCTACCATGCGGCT 1120
 QY 3290 CTTGGCGCTTATACACAGGCGCACTGGGACAACTAAGCGTATCAGAGGTGTCATGCC 3348
 Db 1121 CTTGGCGCGAGGCTTCACTGCGCACTCATCCCTGTCATGAGAGATCCACAGC 1179

RESULT 9

US-09-833-381-427/C
 ; Sequence 427, Application US/09833381
 ; Patent No. 6672186
 ; GENERAL INFORMATION:
 ; APPLICANT: Robison, Keith E.
 ; TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs
 ; FILE REFERENCE: 5800-119
 ; CURRENT APPLICATION NUMBER: US/09/833,381
 ; CURRENT FILING DATE: 2001-04-11
 ; PRIOR APPLICATION NUMBER: 09/516,448
 ; PRIOR FILING DATE: 2000-02-29
 ; NUMBER OF SEQ ID NOS: 2050
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 427
 ; LENGTH: 524
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-09-833-381-427

Query Match 2.9%; Score 110.4; DB 4; Length 524;
 Best Local Similarity 57.5%; Pred. No. 6.8e-19;
 Matches 198; Conservative 0; Mismatches 146; Indels 0; Gaps 0;
 QY 2995 CACATGAAGACTCAGAGTGTGAGAAACCTTCCGCTGTGCGCGTGTCCATAGGCTCT 3054
 Db 512 CACAGGAGGTCACACACAGGAGACAGACCCCTTTCAATGTGATCTCTGTGCGAAGCGCTC 453
 QY 3055 GCTCATCTGGATACCTGAAAGCGCACCGGCTCCACAGAGAGAGAAAGCCCTACAAG 3114
 Db 452 TTGACGCGCTCAGACCTCCGAGTTCACAGGAGTCCACACTGGGAGAGGCGCTCATG 3393
 QY 3115 TGCCCCCTCTGTCCGTATGCTGTGGCAACCTGGCCAACTCAAGCGTGTGTCGATC 3174
 Db 392 TGTGAGCTCTGCCAAAGCGGTTCCGCCACGAGTCCACCTTACAGGGGCCACAGAGGATC 333
 QY 3175 CACTCTGGTGACAAACCTTTTCGGGTAGCTTTGGCACTTACAGCTGCAACCCAGAGTATG 3234
 Db 332 CACACCGGGGAGAGGCGCTTTAAATGTAATATCTGACGAAAGTTTTCAGCCACAGGGG 273
 QY 3235 AACCTCAAACTCATATGCTGCGACACACGCGGCGAGAGCCCTTCCGCTGTGCCACTGC 3294
 Db 272 AACCTGAAGCTTACCAGGCGCACCCACTCCGAGAGAGGCGCTTACAAATGTCCACCTGT 213
 QY 3295 GCCTATACCAGCGCACTGGGACAACTACAGCGTTCATCAGAA 3338
 Db 212 CAAAAGGCGCTTCCGTGACGTGGGGACATTCAAGCGTCACTGTAA 169

RESULT 10

US-09-418-640-3
 ; Sequence 3, Application US/09418640
 ; Patent No. 640125
 ; GENERAL INFORMATION:
 ; APPLICANT: Jennifer K. Taylor
 ; APPLICANT: Lex M. Cowser

; TITLE OF INVENTION: ANTISENSE MODULATION OF BCL-6 EXPRESSION

; FILE REFERENCE: RTS-0102

; CURRENT APPLICATION NUMBER: US/09/418,640

; CURRENT FILING DATE: 1999-10-15

; NUMBER OF SEQ ID NOS: 89

; SEQ ID NO 3

; LENGTH: 3536

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (328)..(2448)

US-09-418-640-3

Query Match 2.6%; Score 100.4; DB 3; Length 3536;

Best Local Similarity 54.3%; Pred. No. 7.4e-16;

Matches 203; Conservative 0; Mismatches 171; Indels 0; Gaps 0;

QY 2920 CAGGCGCTGTGACAAAGGCTTTGCTGTAGTTTATGCCCCCTTTGCCACTCACTACCCC 2979

Db 1942 CAGACCCACAGTGACAAACCCCTACAAGTGTGACCGCTGCCAGGCTCTTCCCGTACAAG 2001

QY 2980 AACCACTCGCTCGGCACATCAAGACTCACAGTGTGAGAAACCTTCCGCTGTGCCGC 3039

Db 2002 GGCACCTCGGCACACAGACCGCTCCATACCGTGTGAAACCTTATCGTTGCAATC 2061

QY 3040 TGTCCATACGCTCTGTCTCATCTGATTAACCTGAACCGCACCGCGCTCCACAGGA 3099

Db 2062 TGTGGGGCCAGTTTCAACCGCGCAGCAACCTGAAACCCACACATCTCAATCTCTGGA 2121

QY 3100 GAAGAGCCCTACAGTGTGCCCCCTCTGTCTCGTATCGCTGTGCAACCTGCGCAACCTCAAG 3159

Db 2122 GGAAGCCCTCAATATGGAACCTGCGAGCCAGATTTGTACAGTGTGGCCACCTCCGT 2181

QY 3160 CTTATGTGCGATCCACTCTGTGTGACAAACCTTTTCGGTGTAGCCCTTTGCAACTACAGC 3219

Db 2182 GCCCATGTGCTTATCCACACTGTGTGAGAAAGCCCTATCCCTGTGAAATCTGTGGCACCCGT 2241

QY 3220 TGCACACAGATGATGAACCTCAAGCTCATATGCTGCGACACACGCGGAGAGCCCTTC 3279

Db 2242 TTCGGCACCTTCAGACTCTGAAGACCCTGCGAATCCACACAGAGAGAAACCTTAC 2301

QY 3280 CGCTGTGCCACCTG 3293

Db 2302 CATTGTGAGAAGTG 2315

RESULT 11

US-08-074-967-1

; Sequence 1, Application US/08074967

; Patent No. 5641672

; GENERAL INFORMATION:

; APPLICANT: Dalia-Favera, Riccardo

; APPLICANT: Chaganti, R.S.K.

; TITLE OF INVENTION: CLONING AND USES OF THE GENETIC LOCUS

; TITLE OF INVENTION: bcl-6

; NUMBER OF SEQUENCES: 2

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Cooper & Dunham

; STREET: 30 Rockefeller Plaza

; CITY: New York

; STATE: New York

; COUNTRY: United States of America

; ZIP: 10112

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent in Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/074,967

; FILING DATE:

; CLASSIFICATION: 435

;

; ATTORNEY/AGENT INFORMATION:

; NAME: White, John P.

; REGISTRATION NUMBER: 28,678

; REFERENCE/DOCKET NUMBER: 0575/43771

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 977-9550

; TELEFAX: (212) 664-0525

; TELEX: 422523 COOP UI

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 3720 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA

; FEATURE:

; NAME/KEY: CDS

; LOCATION: 328..2445

US-08-074-967-1

Query Match

Best Local Similarity 54.3%; Pred. No. 7.6e-16;

Matches 203; Conservative 0; Mismatches 171; Indels 0; Gaps 0;

QY 2920 CAGGCGCTGTGACAAAGGCTTTGCTGTAGTTTATGCCCCCTTTGCCACTCACTACCCC 2979

Db 1942 CAGACCCACAGTGACAAACCCCTACAAGTGTGACCGCTGCCAGGCTCTTCCCGTACAAG 2001

QY 2980 AACCACTCGCTCGGCACATCAAGACTCACAGTGTGAGAAACCTTCCGCTGTGCCGC 3039

Db 2002 GGCACCTCGGCACACAGACCGCTCCATACCGTGTGAAACCTTATCGTTGCAATC 2061

QY 3040 TGTCCATACGCTCTGTCTCATCTGATTAACCTGAACCGCACCGCGCTCCACAGGA 3099

Db 2062 TGTGGGGCCAGTTTCAACCGCGCAGCAACCTGAAACCCACACATCTCAATCTCTGGA 2121

QY 3100 GAAGAGCCCTACAGTGTGCCCCCTCTGTCTCGTATGCTGTGCAACCTGCGCAACCTCAAG 3159

Db 2122 GGAAGCCCTCAATATGGAACCTGCGAGCCAGATTTGTACAGTGTGGCCACCTCCGT 2181

QY 3160 CTTATGTGCGATCCACTCTGTGTGACAAACCTTTTCGGTGTAGCCCTTTGCAACTACAGC 3219

Db 2182 GCCCATGTGCTTATCCACACTGTGTGAGAAAGCCCTATCCCTGTGAAATCTGTGGCACCCGT 2241

QY 3220 TGCACACAGATGATGAACCTCAAGCTCATATGCTGCGACACACGCGGAGAGCCCTTC 3279

Db 2242 TTCGGCACCTTCAGACTCTGAAGACCCTGCGAATCCACACAGAGAGAAACCTTAC 2301

QY 3280 CGCTGTGCCACCTG 3293

Db 2302 CATTGTGAGAAGTG 2315

RESULT 12

US-08-553-541B-1

; Sequence 1, Application US/08553541B

; Patent No. 582858

; GENERAL INFORMATION:

; APPLICANT: Dalia-Favera, Riccardo

; APPLICANT: Chaganti, Raju S.K.

; TITLE OF INVENTION: CLONING AND USES OF THE GENETIC LOCUS

; TITLE OF INVENTION: bcl-6

; NUMBER OF SEQUENCES: 9

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Cooper & Dunham LLP

; STREET: 1185 Avenue of the Americas

; CITY: New York

; STATE: New York

; COUNTRY: United States of America

; ZIP: 10036

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

;

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA: /08/553,541B
FILING DATE: May 28, 1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 0575/43771-A-PCT-US
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0525
TELEX: 422523 COOP UI
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3720 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 328..2445
US-08-553-541B-1

Query Match 2.6%; Score 100.4; DB 2; Length 3720;
Best Local Similarity 54.3%; Pred. No. 7.6e-16;
Matches 203; Conservative 0; Mismatches 171; Indels 0; Gaps 0;
QY 2920 CAGGGCCCTGGTGACAAAGGCTTTGCCCTGTAGTTTATGCCCTTTGCCACTCACTACCCC 2979
DB 1942 CAGACCCACAGTGACAAACCTCAAGTGTGACCGCTGCCAGSCCTCTTCGCTACAAG 2001
QY 2980 AACCACTGCTCGGCACATGACATCAAGTGTGACAAACCTTTGCCACTCACTACCCC 3039
DB 2002 GGCACCTGCGCCAGCCACAGACCGTCCATACCGGTGAGAAACCTTATCGTGCAACATC 2061
QY 3040 TGTCCATACGCTCTGCTCTATCTGTGATAACCTGAAACGGCAGCGCGCTCCACACAGGA 3099
DB 2062 TGTGGGCCCACTTCAACCGCCAGCCCAACCTGAAACCCACACTCGAATTCACCTGGA 2121
QY 3100 GAAAGCCCTACAGTGGCCCTCTGTCGTATGCTGTGGCAACCTTGCCACCTCAAG 3159
DB 2122 GAGAGCCCTTACAAATCGAAACCTGGGAGCCAGATTTGTACAGTGGCCCACTTCGT 2181
QY 3160 CGTCATGCTGCACTCCACTCTGCTGTGACAAACCTTTTGGGTGTAGCCCTTTGCAACTACAGC 3219
DB 2182 GCCCATGTGCTTATCCACACTGGTGTGAGAGCCCTATCCCTGTGAATCTGTGCAACCGT 2241
QY 3220 TGCACACAGATGATGAACCTCAACCTCATATGCTGGGACACACGGGCGAGAGCCCTTC 3279
DB 2242 TTCGGCACCTTCAAGCTCTGAAGAGCCCTCGAATTCACACAGGAGAGAAACCTTAC 2301
QY 3280 CGCTGTGCCACCTG 3293
DB 2302 CATTGTGAGAAGTG 2315

RESULT 13
US-09-268-202-1
Sequence 1, Application US/09268202
Patent No. 6174997
GENERAL INFORMATION:
APPLICANT: Dalla-Favera, Riccardo
APPLICANT: Chaganti, Raju S.K.
TITLE OF INVENTION: CLONING AND USES OF THE GENETIC LOCUS
TITLE OF INVENTION: bcl-6
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York

STATE: New York
COUNTRY: United States of America
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/268,202
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 0575/43771-A-PCT-US-2
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0525
TELEX: 422523 COOP UI
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3720 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 328..2445
US-09-268-202-1

Query Match 2.6%; Score 100.4; DB 3; Length 3720;
Best Local Similarity 54.3%; Pred. No. 7.6e-16;
Matches 203; Conservative 0; Mismatches 171; Indels 0; Gaps 0;
QY 2920 CAGGGCCCTGGTGACAAAGGCTTTGCCCTGTAGTTTATGCCCTTTGCCACTCACTACCCC 2979
DB 1942 CAGACCCACAGTGACAAACCTCAAGTGTGACCGCTGCCAGSCCTCTTCGCTACAAG 2001
QY 2980 AACCACTGCTCGGCACATGACATCAAGTGTGACAAACCTTTGCCACTCACTACCCC 3039
DB 2002 GGCACCTGCGCCAGCCACAGACCGTCCATACCGGTGAGAAACCTTATCGTGCAACATC 2061
QY 3040 TGTCCATACGCTCTGCTCTATCTGTGATAACCTGAAACGGCAGCGCGCTCCACACAGGA 3099
DB 2062 TGTGGGCCCACTTCAACCGCCAGCCCAACCTGAAACCCACACTCGAATTCACCTGGA 2121
QY 3100 GAAAGCCCTTACAGTGGCCCTCTGTCGTATGCTGTGGCAACCTTGCCACCTCAAG 3159
DB 2122 GAGAGCCCTTACAAATCGAAACCTGGGAGCCAGATTTGTACAGTGGCCCACTTCGT 2181
QY 3160 CGTCATGCTGCACTCCACTCTGCTGTGACAAACCTTTTGGGTGTAGCCCTTTGCAACTACAGC 3219
DB 2182 GCCCATGTGCTTATCCACACTGGTGTGAGAGCCCTATCCCTGTGAATCTGTGCAACCGT 2241
QY 3220 TGCACACAGATGATGAACCTCAACCTCATATGCTGGGACACACGGGCGAGAGCCCTTC 3279
DB 2242 TTCGGCACCTTCAAGCTCTGAAGAGCCCTCGAATTCACACAGGAGAGAAACCTTAC 2301
QY 3280 CGCTGTGCCACCTG 3293
DB 2302 CATTGTGAGAAGTG 2315

RESULT 14
PCT-US94-06669-1
Sequence 1, Application PC/TUS9406669
GENERAL INFORMATION:
APPLICANT: Dalla-Favera, Riccardo
APPLICANT: Chaganti, R.S.K.
TITLE OF INVENTION: CLONING AND USES OF THE GENETIC LOCUS
TITLE OF INVENTION: bcl-6

NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10112
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/06669
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/074,967
FILING DATE: 09-JUN-1993
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 43771-A-PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 977-9550
TELEFAX: (212) 684-0525
TELEX: 422523 COOP UI
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3720 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 328..2445
PCT-US94-06669-1

Query Match 2.6%; Score 100.4; DB 5; Length 3720;
Best Local Similarity 54.3%; Pred. No. 7.6e-16;
Matches 203; Conservative 0; Mismatches 171; Indels 0; Gaps 0;
QY 2920 CAGGCGCTGTGACAAAGGCTTTGCTGTAGTTTATGCCCCCTTGGCACTCACTACCCC 2979
DB 1942 CAGACCCACAGTGACAAACCCTACAAGTGTGACCGCTGCCAGGCTCTCTCCGCTACAAG 2001
QY 2980 AACCACTGGTCCGCACATGAAGACTCAGAGTGTGAGAAACCTTCCGCTGTGCCCGC 3039
DB 2002 GGCACCTCGCCAGCCACAGACCGTCCATACCGGTGAGAAACCTTATCGTTGCCAACATC 2061
QY 3040 TGTCCATACGCTCTGTCTCATCTGGATAACCTGAAACGGCAGCAGCGCTCCACAGGA 3099
DB 2062 TGTGGGGCCAGTTTCAACCGGCCAGCCAACTGAAACCCACACTCGAATTCACCTGGA 2121
QY 3100 GAAAGCCCTCAGAGTGGCCCTCTGTCCGTATGCTGTGCACTGGCACTGGCCACCTCAAG 3159
DB 2122 GAGAAAGCCCTCAGATGGAACCTGGGAGCCAGATTGTACAGTGGGCCACCTCCCGT 2181
QY 3160 CGTATGCTCGCATCCACTCTGSGTGAACAACTTTTCGGTGTAGCCCTTTTGCACACTACAGC 3219
DB 2182 GCCCATGTGCTTATCCACACTGGTGAAGGCCCTATCCCTGTGAAATCTGTGGCACCCGT 2241
QY 3220 TGCACACAGATATGAACCTCAGACGTATATGCTGCGACACAGCGCGGAGAGCCCTTC 3279
DB 2242 TTCGGCACCTTCAGACTCTGAGAGCCCACTGCGAATCCACACAGAGAGAACCTTAC 2301
QY 3280 CGCTGTGCCACTG 3293
DB 2302 CATTGTGAGAAGTG 2315

RESULT 15
US-09-620-312D-309
; Sequence 309, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yundong
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; FILE REFERENCE: Polypeptides
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: dt_FL_genes Version 1.0
; SEQ ID NO 309
; LENGTH: 2789
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (248)..(2659)
US-09-620-312D-309

Query Match 2.5%; Score 96.4; DB 4; Length 2769;
Best Local Similarity 55.8%; Pred. No. 7.1e-15;
Matches 184; Conservative 0; Mismatches 146; Indels 0; Gaps 0;
QY 2984 ACCTGGCTCGGCACATGAAGACTCAGAGTGGTGAGAAACCTTCCGCTGTGCCCGCTGTC 3043
DB 1626 ACCTGAAGGCCACCTGAAGATCCACATCGCTGACGGGCCCTCAAGTCCGAGAGTGTG 1685
QY 3044 CATAGGCTCTGTCTCATCTCGATTAACCTGAACGGCACAGCGCTCCACACAGGAGAA 3103
DB 1686 GGAAGCAGTTTCAACCACTCAGGGAACCTGAAGCGGACCTTCGGATCCACAGCGGGAGA 1745
QY 3104 AGCCCTCAAGTGGCCCTCTGTCCGTATGCTGTGGCAACCTGGCCAACTCAAGCGTC 3163
DB 1746 AGCCCTACGTGTGATCCACTGCGACGACAGTTTGACAGACCCCGGCTCTGCAGCGGC 1805
QY 3184 ATGTCTGATCCATCTCTGGTGAACAACTTTTCGGTGTAGCTTTTGCACACTACAGTGCA 3223
DB 1806 ACGTCCGCAATTCACACAGGTGAGAGCCATGCCAGTGTGTGATGTGCGGTAAAGCCCTCA 1865
QY 3224 ACCAGAGTATGAACCTCAAACTCATATGCTGGACACAGCGCGGAGAGCCCTTCCGCT 3283
DB 1866 CCCAGGCCAGCTCCCTCATCGCCCACTGAGCCACACACCGGGGAGAGCCCTACGCTCT 3283
QY 3284 GTGCCACTCTGGCCCTATACACAGGCCACT 3313
DB 1926 GCGAGCGCTGCGGCAAGAGATTGCTCCAGT 1955

Search completed: May 14, 2004, 05:50:07
Job time : 266 secs

us-09-673-994a-7.rni

Fri May 14 09:10:45 2004


```

; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20040033942A1 870100CB1
US-10-221-625-121

Query Match      38.9%; Score 1498; DB 13; Length 1975;
Best Local Similarity 88.6%; Pred. No. 0;
Matches 1707; Conservative

QY 1917 AGTGGTACTGAGATTCCTTCGACGAAGGTCTCTGGGGCCCTGGTGTGGAGAGCATTT 1976
Db 86 AGTGGTACTGAGATTCCTTCGACGAAGGTCTCTGGGGCCCTGGTGTGGAGAGCATTT 145

QY 1977 GCTACTAGGCCAAGATCTGGAGTTTGAAGAGGAGAGAGAGAGATGAAGGTGACGGCCA 2036
Db 146 GCTACTAGGCCAAGATCTGGAGTTTGAAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 205

QY 2037 CAACGACACAGTCTATGGGCTTTGAGAGAGACTCTGAGAGAGACTCTCAGGGGCCAGACC 2096
Db 206 CAGTGACACAGTCTATGGGCTTCGAGAGAGACTCTGAGAGAGACTCTCTGGGGGCCAGGCC 265

QY 2097 TGGACTTCCCTATGGGCTGAGTACGACGAGTCTTGGGGGGGGCGCGCACTAAGTGGCGA 2156
Db 266 TGGGCTTCCCTATGGGCTGAGGACGATGAGTCTTGGGGGGGGCGCGCACTAAGTGGCGA 325

QY 2157 GAGTGAAGTTGAGGAAACAGGACGAGGCTCAGGGGAGGCGCAGGGGTGAGAGGCCAGCCC 2216
Db 326 GAGTGAAGTTGAGGAGGACGAGGCTCAGGGGAGGCGCAGGGGTGAGAGGCCAGGCC 385

QY 2217 AGCTGTGACGTGTGTGGGGGGCGACAGGTGAGGGGCGGTGTGTGGGGCAGGAGGGCG 2276
Db 386 AGCTGTGACGTGTGTGGGGGGCGACAGGTGAGGGGCGGTGTGTGGGGCAGGAGGGCG 445

QY 2277 GGGTGGGGGGCGCGCGTGTGCCCCCAGCGTTACTGTACTGTGACGCGCTTGTGACCTTCGT 2336
Db 446 GGGTGGGGGGCGCGCGTGTGCCCCCAGCGTTACTGTACTGTGACGCGCTTGTGACCTTCGT 505

QY 2337 GTCCCACTACTCGACCACTGTGAGGCGCACATGACAGACACACAGCGGGAGAGCGGTT 2396
Db 506 GTCCCACTACTCGACCACTGTGAGGCGCACATGACAGACACACAGCGGGAGAGCGGTT 565

QY 2397 CGCTGTGGCGCTGCCATAGCNCAGCCAGTTCGTCAACCTGACGCGACATACCCG 2456
Db 566 CGCTGTGGCGCTGCCATAGCNCAGCCAGTTCGTCAACCTGACGCGACATACCCG 625

QY 2457 CACCACTACTGCGGAGAGCCCTACGGTGTGCCCCCAGCTGCCCCCTTGGCTGCGAGAGCT 2516
Db 626 CACCACTACTGCGGAGAGCCCTACGGTGTGCCCCCAGCTGCCCCCTTGGCTGCGAGAGCT 685

QY 2517 GGGCAACCTGAGCGGCATCAGCGCACCCACACAGCGGCTCCCACTCCTCCTCGCCCAAC 2576
Db 686 GGGCAACCTGAGCGGCATCAGCGTACCCACAGCGGGCCCCCACTCCTCCTCGCCGAC 745

QY 2577 CTGTGGCTTTCAGTGTGCTTCCAGCAACCCGGCTCCAGTCCCAAGAGAGAGG 2636
Db 746 CTGTGGCTTTCAGTGTGCTTCCAGCAACCCGGCTCCAGTCCCAAGAGAGAGG 805

QY 2637 GGGGCAATGCCCCGACGATCAGAAATGCGGTGATCTCTGCGCAGACTTGAATTCATGT 2696
Db 806 GGGGCGGTGCCCCGCGGACCTGAGATGCTCTGCTCTCCAGATTGAGGCTCCATGT 865

QY 2697 GCCACAGGTGTGCGAGTTTCTGCGCAGACTGTGGGCGAGTGTGGGGTGAAGGGGAG 2756
Db 866 GCCACAGGTGTGCGAGTTTCTGCGCAGACTGTGGGCGAGTGTGGGGTGAAGGGGAGG 925

QY 2757 CTGTGTGAATGAGTCCGACCACTGCGAGGCTACTGTTCCTTGGACCTGCGGGG 2816
Db 926 CTGTGTGGAATGAGTCCGACCACTGCGAGGCTACTGTTCCTTGGACCTGCGGGG 985

QY 2817 CTGTGACAGGAACTGGAGAGGGTGAAGGCGACAGGCTGGAGCTGCCATGTGGGGG 2876
Db 986 CTGTGACAGGAACTGGAGAGGGTGAAGGCGTGTGCGGCTGGAGCTGCCATGTGGGGG 1045

```

RESULT 3
US-09-764-991-7625
; Sequence 7625, Application US/09764891

```

QY 2877 CTGCATCGAGGAGAGGCTGGAGGGGTTCACACTGGGGGACCCCGAGGGCCCTGTGTGACAA 2936
Db 1046 CTGCATCGAGGAGAGGCTGGAGGGGTTCAGTGGGGGGCCCGAGGGCCCGCATGTGACAA 1105

QY 2937 AGGCTTTGCCCTGTAGTTTATGCCCTTTTGCACACTCACTATACCCCAACCACTTGGGTGCGCA 2996
Db 1106 AGGCTTTGCCCTGTAGCTTCTGCCCTTTTGCACACTCACTATCCCAACCACTTGGGTGCGCA 1165

QY 2997 CATGAAGACTCACAGTGTGTGAAGAACCTTCGGTGTGCGCGCTGTCCATACGCTCTGCG 3056
Db 1166 CATGAAGACACACAGTGTGTGAAGAACCTTCGGTGTGCGCGCTGTCCATACGCTCTGCG 1225

QY 3057 TCATCTGTGATTAACCTGAAACGSCACAGCGGTCCACACAGAGAGAAAGCCCTACAACTG 3116
Db 1226 TCATCTGTGATTAACCTGAAACGSCACAGCGGTCCATACAGAGAGAGAGCCCTACAACTG 1285

QY 3117 CCCCCCTGTTCGATATGCTGTGGCAACCTGGCCAACTCAAGCGTCAATGTGTCGATCCA 3176
Db 1286 CCCCCCTGTTCGATATGCTGTGGCAACCTGGCCAACTCAAGCGTCAATGTGTCGATCCA 1345

QY 3177 CTCTGGTGAACAACTTTTTCGGTGTAGCTTTTGCACACTACAGCTGCAACCAAGAGTATGAA 3236
Db 1346 CTCTGGTGAACAACTTTTTCGGTGTAGCTTTTGCACACTACAGCTGCAACCAAGAGTATGAA 1405

QY 3237 CCTCAAACTCATATGCTGCGCACACACGGGGGAGAGCCCTTCCGCTGTGCGCACTTGGCGC 3296
Db 1406 CCTCAAACTCATATGCTGCGCACACACGGGGGAGAGCCCTTCCGCTGTGCGCACTTGGCGC 1465

QY 3297 CTATACACAGGCCACTGGGACAACTACAGCGTCTATCAGAGGTGCAAGGCCATGTGGTGG 3356
Db 1466 CTATACACAGGCCACTGGGACAACTACAGCGTCTATCAGAGGTGCAAGGCCATGTGGTGG 1525

QY 3357 AGCAGAGAGCGCTGTGCTCTCTGCCCTCTGAGGGTGGGGCCCCACCTCATAGGCCACCTC 3416
Db 1526 AGCAGAGAGCGCTGTGCTCTCTGCCCTCTGAGGGTGGGGCCCCACCTCATAGGCCACCTC 1585

QY 3417 TGTGTTGAGCACTCGGGGTTCAGAGCCCTTGGGTGTACTGTGTAGCAGGCTTTTCAATTC 3476
Db 1586 TGTGTTGAGCACTCGGGGTTCAGAGCCCTTGGGTGTACTGTGTAGCAGGCTTTTCAATTC 1645

QY 3477 AGACTCACTTGAACCTAACTAGGTTCCTTTTACCTGGGGCTCTAGGAATTAGCCCTATGCC 3536
Db 1646 AGACTCACTTGAACCTAACTAGGTTCCTTTTACCTGGGGCTCTAGGAATTAGCCCTATGCC 1672

QY 3537 TCCTGCAATTTTATACAAATGAACTAGAAACCACTTTTCCCTTCTCCCGCTGTGTGAG 3596
Db 1673 TCCTGCAATTTTATACAAATGAACTAGAAACCACTTTTCCCTTCTCCCGCTGTGTGAG 1728

QY 3597 GGCTCCACACAGACTAACCTTAGGCACTATATGACCGAGCTGTAATCCCATGGTTCAGGGG 3656
Db 1729 GGCTCCACACAGACTAACCTTAGGCACTATATGACCGAGCTGTAATCCCATGGTTCAGGGG 1788

QY 3657 -CCATATAGACAGGGGA-CTTGTCTTAGCTACGTACCGATGAGCTAGTGAATGATGGG 3714
Db 1789 CCATATAGACAGGGGACCTTGTCTTGTAGCTAGGCACTTACAGAGCTCAGTGAAGAGG 1848

QY 3715 CTTTGGATTACCGCCACTGCTCCAGAGGCTATGATGAACCTGTTGGGAG-CTGGCCCA 3773
Db 1849 CTTTGGATTACCGCCACTGCTCCAGAGGCTATGATGAACCTGTTGGGAG-CTGGCCCA 1908

QY 3774 GCCTTTTAC----TGTTTAACTTATTTTCACTGCTTTTATATTAAGGAAACACTTAA 3829
Db 1909 GCCTTTTAC----TGTTTAACTTATTTTCACTGCTTTTATATTAAGGAAACACTTAA 1968

QY 3830 AAAAAA 3836
Db 1969 AAAAAA 1975

```

```

; Publication No. US20030077808A1
;
; GENERAL INFORMATION:
;
; APPLICANT: Rosen et al.
;
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;
; FILE REFERENCE: PCC06
;
; CURRENT APPLICATION NUMBER: US/09/764,891
;
; CURRENT FILING DATE: 2001-01-17
;
; Prior application data removed - consult PALM or file wrapper
;
; NUMBER OF SEQ ID NOS: 10231
;
; SOFTWARE: PatentIn Ver. 2.0
;
; SEQ ID NO 7625
;
; LENGTH: 9266
;
; TYPE: DNA
;
; ORGANISM: Homo sapiens
;
; US-09-764-891-7625

```

Query Match	31.1%	Score 1195;	DB 10;	Length 9266;	
Best Local Similarity	81.3%;	Pred. No. 0;			
Matches 1568;	Conservative	0;	Mismatches 216;	Indels 144;	Gaps 10;
QY	2003	AAGAGGAAGAGGAGGATGAAGGTGACGGGCACCAACACGACGAGCTCATGGCTTTTGAGA	2062		
DB	1441	AACTGAAAAGGGGAGATTGGAGCGCTGATCCCTACCTATGCTGATGTCTCTCTTATGT	1500		
QY	2063	GAGACTCTGAAGAGAGACTCTCAGGGGGCCAGACTCGGATTCCTCCCTATGGCTGAGTGACG	2122		
DB	1501	CTATTTTACCAGGAGACTCTCTGGGGGCCAGGCGCTGGGCTTTCCCTATGGCTTGACGACG	1560		
QY	2123	ACGAGTCTGGGGGGCGCGGCCACTAAGTTCGGAGAGTGAAGTTTGAGAACCAACGACGAGG	2182		
DB	1561	ATGAGTCTGGGGGGCGCGGCCACTAAGTTCGGAGAGTGAAGTTTGAGGAGCCAGCCAGCG	1620		
QY	2183	GTCAGGGGAGCCAGGGGTGAGNGGCCAGGCCAGCCCTGTCTGAGCTGTGTGGGGGGCCGA	2242		
DB	1621	GTCAGGGGAGCCAGGGGTGAGAGGCCAGGCCAGCCCTGCCAGCTGTGTGGGGGGCCGA	1680		
QY	2243	CAGGTGAGGGGGCGGTGTGTGGGGCCAGAGGGGGGGTGGGGGGCCCCGCTGCCCCCAAC	2302		
DB	1681	CAGGTGAGGGGGCGGTGTGTGGGGCCAGAGGGGGGGTGGGGGGCCCCGCTGCCCCCAAC	1739		
QY	2303	GGTTACTGTACTCATGCGGGGTGCGCTTTTCGTGCCCACTACTCTGAGCCACTCTGAAGC	2362		
DB	1740	GGCTACTGTACTCATGCGGGGTGCGCTTTTCGTGCCCACTACTCTGAGCCACTCTGAAGC	1799		
QY	2363	GGCAATGTCAGACACACAGCGGGGAGAACGGCTTCCGCTGTGGCCGCTGCCCATACGCT	2422		
DB	1800	GGCAATGTCAGACACACAGCGGGGAGAACGGCTTCCGCTGTGGCCGCTGCCCATACGCT	1859		
QY	2423	CAGCCCACTTCGTCAAACCTGACGGACATACCCGACCCCATATCTGGCGAGAACCTTACC	2482		
DB	1860	CAGCCCACTTCGTCAAACCTGACCGACATACCCGACCCCATATCTGGCGAGAACCTTACC	1919		
QY	2483	GTTGTCCCAATGCGCCCTTTGCCCTGACAGAGCTTGGCGAACCTGAGCGGCACTCAGGCA	2542		
DB	1920	GCTGTCCCAATGCGCCCTTTGCCCTGACAGAGCTTGGCGAACCTGAGCGGCACTCAGGCA	1979		
QY	2543	CCCAACAGGGGCTCCCACTCCTCCCTGCCCAACCTGTGGGCTTTTCGATGTCTCTCCAC	2602		
DB	1980	CCCAACAGGGGCTCCCACTCCTCCCTGCCCAACCTGTGGGCTTTTCGATGTCTCTCCAC	2039		
QY	2603	GACCAACCGGGCTCCAGTCCCAAGAGGAGGGGACAAATGCCCG	2651		
DB	2040	GACCAACCGGGCTCCAGTCCCAAGAGGAGGGGACAAATGCCCG	2099		
QY	2652	-----	2651		
DB	2100	TAAGACACACAGGGACCAAGATCTTGGGACATGGGTGGCTGACCCCTAGGAATGCTTGG	2159		
QY	2652	-----ACGATCAGAAATGCGCTGATCTCCCTGCCAGACTTGAGT	2688		
DB	2160	ATTGGATTCATAGCCAGGTCTTTGTCCCAACAGATGCTCTCTCTCCAGATTTGAGC	2219		
QY	2689	CTTCATGTGCCACCAAGGTGGTGCCAGTTTCTCTGCCAGACTGTGGGAGCTCGGGGGTGAA	2748		

Db	2220	CTCCATGTGCCACAGGT - GTGCCAGTTTCTGCCAGACTGTGGCAGCTCGGGGTGA	2278
Qy	2749	GGGAGAGCTTCTGTGGAACTGGATCCGAAACA	2808
Db	2279	GGGAGGGCTCTGGGAGCTGGATCAACA	2338
Qy	2809	TGCGGGGTGTGGACAGAACTGGAGAGGGTGAAGGACAGAGCTGGGAGCTGCCATG	2868
Db	2339	TGCGGGGTGTGGACAGAGCTGGAGAGGGTGAAGGTAGTCGGCTGGGAGCTGCCATG	2398
Qy	2869	TGTGGCGCTGTGATCGAGAGAGAGCTGGAGGGTGTGCACTGGGGGACCCAGGGGCGCT	2928
Db	2399	TGTGGCGCTGTGATCGAGAGAGAGCTGGAGGGGTGCAGTGGGGGCCACACGCTGT	2458
Qy	2929	GGTGACAAAGGCTTTCCTGTATGTATGTGCCCTTTTGGCACTCACTACCCCAACCACTG	2988
Db	2459	-----ACAAAGGCTTTCCTGTATGTATGTGCCCTTTCGCACTCACTATCCCAACCACTG	2514
Qy	2989	GCTCGGCATGAAGACTCACAGTGGTGAGAAACCTTTCGGGTGTGCCGTGTCCATAC	3048
Db	2515	GCCCGGCATGAAGACACACAGTGTGAAAGCCCTTTCGGGTGTGCCGTGTCCAT	2574
Qy	3049	GCCTCTGCTCATCTGGATAACCTGAAACGGCACAGCGGTGCCACAGGAGAAAGGCC	3108
Db	2575	GCCTCTGCTCATCTGGATAACCTGAAACGGCACAGCGGTGCCATACAGGAGAAAGGCC	2634
Qy	3109	TACAGTGGCCCTCTGTCCGTATGCTGTGGCAACCTGGCCAACTCAAGCGTCTATGCT	3168
Db	2635	TACAGTGGCCCTCTGTCCCTTATGCTGTGGCACTGTGGCCAACTCAAGCGTCTATGCT	2694
Qy	3169	CGCATCACTCTGCTGACAAACCTTTCGGGTGTAGCCTTTTGCAACTACAGCTGCACACAG	3228
Db	2695	CGCATCACTCTGCTGACAAACCTTTCGGGTGTAGCCTTTTGCAACTACAGCTGCACACAG	2754
Qy	3229	AGTATGAACCTCAAAAGTCTATGTCTGACACACAGGGCGAGAGCCCTTCGCTGTGCC	3288
Db	2755	AGCATGAACCTCAAAAGTCTATGTCTGCGCACACAGGGCGAGAGCCCTTCGCTGTGCC	2814
Qy	3289	ACCTGCGCTTATACACAGGCCACTGGGACAACTACAAGCGTCACTCAGAAAGTGCATGCG	3348
Db	2815	ACCTGCGCTTATACACAGGCCACTGGGACAACTACAAGGCCACTCAGAAAGTGCATGCG	2874
Qy	3349	CATGTTGGAGCAGAGGGCTGTGCTCTCTGCCCCGTGAGGCTGGGCCCCCACTCATAGC	3408
Db	2875	CACGGTGGGGCAGAGGGCTGTGCTCTCTGCCCCGTGAGGCTGGGCCCCCACTCATAGC	2934
Qy	3409	CCACCTCTGTTTGTAGCACTCGGGTCCAGCAGCCCTGGGTGCTACTGCTAGCAGGGCT	3468
Db	2935	CCACCTCTGTTTGTAGCACTCTGGGGCCCAACAGCCCTGGGAGTGTGTGACCGCGGCT	2994
Qy	3469	CTTCATTTCAGACTCACCTTGAACCTAAGTTCCTTTTACCTGGGGCTCTAGGAATTAGC	3528
Db	2995	GTCCACACAGACTCATCTGAAC-----TAGG 3021	
Qy	3529	CCTATGCTCTGTGATTTTATACAAATGAATAGAAACACACCTTTCCTTCTCCCGCCG	3588
Db	3022	TCTTCTTCCCCTATGTTTATACAGCGGACCAAGAGCCACCTT-----TTTCTCCCGCCG	3077
Qy	3589	TGTTTCAGGGGTCCACACAGACTAACCTAGGCATATATGGACAGCGCTGAATCCCATGG	3648
Db	3078	TGGCCAGGGGTCCACACAGACTAACCTAGGCATATATAGGACACAGCCCAACCCCATGGG	3137
Qy	3649	TCAGGGGG-CCATATAGACAGGGGA-CTTGTCTTAGTCACTACGTACAGATGAGCTAAGT	3706
Db	3138	CGGGGGGGCCCATATGGAACACAGGGGACCTTTCCTTGACTGAGGCGCTTCACGAGCTCAGT	3197
Qy	3707	GATTAGGGCTTGTGATTCACCGCCACCTGCTCCACAGAGGCTATGGATGAACCTGTGTGGAG	3766
Db	3198	GAGAAGGGCCCTGTATTACCTCCACTGCCCCCAGGGGCTGTGCACAAACCGGCTGGGG	3257
Qy	3767	-CTGCCCGAGCCCTTTTAC-----TGTTTAACTATTTCAGTGTCTTTTAAATAAGGAACA	3821

Db 3258 ACTGCCAGGCTCCACCTGTTATTATTAACCTATTTCAGTGTCTTTATTAATAAGAAACA 3317
QY 3822 CTAACAGA 3829
Db 3318 CTAACAAA 3325

RESULT 4

US-10-094-749-1126

; Sequence 1126, Application US/10094749

; Publication No. US20030219741A1

; GENERAL INFORMATION:

; APPLICANT: ISOGAI, TAKAO

; APPLICANT: SUGIYAMA, TOMOYASU

; APPLICANT: OTSUKI, TETSUJI

; APPLICANT: WAKAMATSU, AI

; APPLICANT: SATO, HIROYUKI

; APPLICANT: ISHII, SHIZUKO

; APPLICANT: YAMAMOTO, JUN-ICHI

; APPLICANT: ISONO, YUUKO

; APPLICANT: HIO, YURI

; APPLICANT: OTSUKA, KAORU

; APPLICANT: NAGAI, KEIICHI

; APPLICANT: TRIE, RYOTARO

; APPLICANT: TAMECHIKI, ICHIRO

; APPLICANT: SEKI, NAOHICO

; APPLICANT: YOSHIKAWA, TSUTOMU

; APPLICANT: OTSUKA, MOTOYUKI

; APPLICANT: NAGAHARI, KENJI

; APPLICANT: MASUHO, YASUHIKO

; TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA

; FILE REFERENCE: 084335/0160

; CURRENT APPLICATION NUMBER: US/10/094, 749

; CURRENT FILING DATE: 2002-03-12

; PRIOR APPLICATION NUMBER: 60/350,435

; PRIOR FILING DATE: 2002-01-24

; PRIOR APPLICATION NUMBER: JP 2001-328381

; PRIOR FILING DATE: 2001-09-14

; NUMBER OF SEQ ID NOS: 3381

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 1126

; LENGTH: 1456

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-094-749-1126

Query Match 29.0%; Score 1115.2; DB 16; Length 1456;

Best Local Similarity 87.4%; Pred.No. 0;

Matches 1305; Conservative 0; Mismatches 144; Indels 44; Gaps 6;

QY 2341 CACTACTCGAGCCACCTGAAGCGGCACATGCAACACACAGCGGGAGAGCCGTTCCGC 2400

Db 1 CACTACTCGAGCCACCTGAAGCGGCACATGCAACACACAGCGGGAGAGCCGTTCCGC 60

QY 2401 TGTGGCGGTGCCCATACGNCATGCGCCAGTTGCTACACCTGACGCGACATACCGGACC 2460

Db 61 TGTGGCGGTGCCCATACGNCATGCGCCAGTTGCTACACCTGACGCGACATACCGGACC 120

QY 2461 CATACTGGCGAGAGCCCTACCGTTGTCCGCCACTGCCCCCTTTGCGCTGCGAGCAGCCTGGGC 2520

Db 121 CACACTGGCGAGAGCCCTACCGTTGTCCGCCACTGCCCCCTTTGCGCTGCGAGCAGCCTGGGC 180

QY 2521 AACCTGAGGGGATCAGCGCACCCACAGAGGCTCCGACTCCCTCCCTGCGCAACCTGT 2580

Db 181 AACCTGAGGGGATCAGCGGTACCCACGAGGCCCCCCTCCCTCCCTGCGCAACCTGT 240

QY 2581 GGCTTTTCGATGTGTCTCCAGACCAACCCCGGCTCCGAGTCCACAGAGCAGAGGGGG 2640

Db 241 GGCTTTTCGATGTGTCTCCAGACCAACCCCGGCTCCGAGTCCACAGAGCAGAGGGGG 300

QY 2641 ACAAATGCCCGGAGATCAGAAAATGGCTGATCCCTGCGCAGATTCAGTCTTCATGTGCCA 2700

Db 301 GCGGTGCCCCGGGACCTGAAGATGCTGTCTCTTCCAGATTTTGAAGCTCCATGTGCCA 360

3454	ACTGGTAGCAGGGCTCTTCATTACAGACTCACCTTGAACTAACTAGGTTCCTTTTACCTGGG	3513
Db		1241
1204	GCTGGCAGCCGGGCTGTCCACACAGACTCATCTCGAC	
3514	GCTCTAGGAATTAGCCCTATGCGCTCTGCAATTTTATACAAATGAACCTAGGCACTATATGGACCA	3573
Db	-----TAGGTCCTCTCTTCCCACTGTTTTATACAGACGGACCAAGACCCACCTT-	1289
3574	CCCTTTCTCCCCGCTGCTCAGGGGCTCCACACAGACTAACTAGGCACTATATGGACCA	3633
Db	---TTTCTCCCCGCTGCGCAGGGGCTCCACACAGACTAACTAGGCACTATATGGACCA	1346
1290	GCCTGAATCCCATGGTCCAGGGG-CCATATAGACACAGGGGA-CTTGCTTTAGCTCAAGTA	3691
3634	GCCCAACCCCATGGCGGGGGCCCATATGACACAGGGGACCTTGCTTGACTGAGGCA	1406
Db		
1347	CCAGATGAGCTAAGTGATTTAGGGCCTTGCAATTCACCGCCACTGCTCCACAGAGGCTATGGA	3751
3692	CTTCACGAGCTCAGTGAGAAGGGCCCTGTATTCACTCCACTGCCCCCAGGGGCTGTGGA	1466
Db		
1407	TGAACCTGGTGGAG-CTGCGCCAGCCTTTTAC-----TGTTTTAACTTATTTCAGTGCTTT	3806
3752	CACACCGGCTGGGGGACTGCCAGGCTCCACCTGTTTATTTAACTTATTTCAGTGCTTT	1526
Db		
1467	ATAATAAGGAAACACTAAACAGAAAAA	3838
3807	ATAATAAGGAAACACTAAACAGAAAAA	1558
Db		

RESULT 7

US-09-984-245-48

; Sequence 48, Application US/09984245

; Patent No. US20020165374A1

; GENERAL INFORMATION:

; APPLICANT: Young et al.

; TITLE OF INVENTION: 87 Human Secreted Proteins

; FILE REFERENCE: PZ004P1

; CURRENT APPLICATION NUMBER: US/09/984,245

; CURRENT FILING DATE: 2001-10-29

; PRIOR APPLICATION NUMBER: 09/154,707

; PRIOR FILING DATE: 1998-09-17

; PRIOR APPLICATION NUMBER: PCT/US98/05311

; PRIOR FILING DATE: 1998-03-19

; PRIOR APPLICATION NUMBER: US 60/041,277

; PRIOR FILING DATE: 1997-03-21

; PRIOR APPLICATION NUMBER: US 60/042,344

; PRIOR FILING DATE: 1997-03-21

; PRIOR APPLICATION NUMBER: US 60/041,276

; PRIOR FILING DATE: 1997-03-21

; PRIOR APPLICATION NUMBER: US 60/041,281

; PRIOR FILING DATE: 1997-03-21

; PRIOR APPLICATION NUMBER: US 60/048,094

; PRIOR FILING DATE: 1997-05-30

; PRIOR APPLICATION NUMBER: US 60/048,350

; PRIOR FILING DATE: 1997-05-30

; PRIOR APPLICATION NUMBER: US 60/048,188

; PRIOR FILING DATE: 1997-05-30

; PRIOR APPLICATION NUMBER: US 60/048,135

; PRIOR FILING DATE: 1997-05-30

; PRIOR APPLICATION NUMBER: US 60/050,937

; PRIOR FILING DATE: 1997-05-30

; PRIOR APPLICATION NUMBER: US 60/048,187

; PRIOR FILING DATE: 1997-05-30

; PRIOR APPLICATION NUMBER: US 60/048,099

; PRIOR FILING DATE: 1997-05-30

; PRIOR APPLICATION NUMBER: US 60/048,352

; PRIOR FILING DATE: 1997-05-30

; PRIOR APPLICATION NUMBER: US 60/048,186

; PRIOR FILING DATE: 1997-05-30

; PRIOR APPLICATION NUMBER: US 60/048,069

; PRIOR FILING DATE: 1997-05-30

; PRIOR APPLICATION NUMBER: US 60/048,095

```

; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,131
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,096
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,355
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,160
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,351
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,154
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/054,804
; PRIOR FILING DATE: 1997-08-05
; PRIOR APPLICATION NUMBER: US 60/056,370
; PRIOR FILING DATE: 1997-08-19
; PRIOR APPLICATION NUMBER: US 60/060,862
; PRIOR FILING DATE: 1997-10-02
; NUMBER OF SEQ ID NOS: 343
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 48
; LENGTH: 851
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-984-245-48

Query Match      13.7%; Score 528.4; DB 9; Length 851;
Best Local Similarity 83.3%; Pred. No. 3.6e-146;
Matches 701; Conservative 0; Mismatches 96; Indels 45; Gaps 7;

QY 2995 CACATGAAGACTCACAGTGTGTGAGAAACCTTCGCTGTGCGCGTGTCCATACGCTCT 3054
DB 1 CACATGAAGACACACAGTGTGTGAGAACCTTCGCTGTGCGCGTGTCTTATGCTCT 60

QY 3055 GCTCATCTGATTAACCTGAAACGGCACAGCGGTCCACACAGAGAGAAAGCCCTACAAG 3114
DB 61 CCTCATCTGATTAACCTGAAACGGCACAGCGGTCCATACAGAGAGAGAGCCCTACAAG 120

QY 3115 TGCCCCCTCTGTCCGTATGCTGTGGCAACCTGGCCAACTCAAGCGTATGTGTCGATC 3174
DB 121 TGCCCCCTCTGTCCGTATGCTGTGGCAACTGTGGCAACCTCAAGCGTATGTGTCGATC 180

QY 3175 CACTCTGTGACAAACCTTTTCGCTGTAGCTTTGCACTACAGCTGCAACACAGAGATG 3234
DB 181 CACTCTGTGACAAACCTTTTCGCTGTAGCTTTGCACTACAGCTGCAACACAGAGATG 240

QY 3235 AACCTCAAAGCTCATATGCTGCGACACACGCGGAGAGAGCCCTTCGCTGTGCCACTGC 3294
DB 241 AACCTCAAAGCTCATATGCTGCGACACACGCGGAGAGAGCC-TTCCGCTGTGCCACTGC 299

QY 3295 GCCTATACCAAGGCCACTGGCAACTACAAGCGTCAAGAGGTGCAATGGCCATGGT 3354
DB 300 GCCTATACCAAGGCCACTGGCAACTACAAGCGCCACAGAGGTGCAATGGCCATGGT 359

QY 3355 GGAGCAGAGGCGCTGTGTCTCTGCGCCCTAGGGCTGGGCGCCCACTCATAGCCCAACC 3414
DB 360 GGAGCAGAGGCGCTGTGTCTCTGCGCTGTAGGGCTGGGCGCCCACTCATAGCCCAACC 419

QY 3415 TCTGTTTGTAGACTCGGGTTCAGAGCCCTGGGTGTACTGTGTAGCAGGCTCTTCAT 3474
DB 420 TCTGTTTGTAGCTTCGGGGGCCACAGCCCTGGGGAGCTGTGGCAGCGGGCTGTCCAC 479

QY 3475 TCAGACTCACCTTGAATTAACCTAGGTCTTTTACCTGGGGCTCTAGGAATTAGCCCTATG 3534
DB 480 ACAGACTCATCTGAATAGGTCTTCTTC----- 509

QY 3535 CCTCTGCAATTTATACAAATGAATAGACTAGAACACCTTCCCTTTCTCCCGCTGTGCA 3594
DB 510 ---CCCATGTTTATACAGCGGACCAAGACCTT-----TTTCTCCCGCTGTGCCA 562

QY 3595 GGGGCTCCACACAGACTAACCTAGGCACTATATGACAGCCCTCAATCCCATGCTCAGGG 3654
DB 511 GGGGCTCCACACAGACTAACCTAGGCACTATATGACAGCCCTCAATCCCATGCTCAGGG 3654

```

```

Db 563 GGGGCTCCACACAGACTAACGTAGGCATATAGAGACAGCCCAACCCCATGGGGGGG 622
QY 3655 GG-CCATATAGACACAGGGA-CTTGTCTTAGCTCAGTACACAGATGAGCTAAGTATTAG 3712
Db 623 GGCCCATATGGACACAGGACCTTGCCTTGACTGAGGCACTTCAAGGCTCAGTGAGAAG 682
QY 3713 GGCCTTGGATTACCGCCACTGCTCCAGAGGCTATGGATGAACCTGTTGGGAG-CTGCC 3771
Db 683 GGGCCTGTATTACCTCCACTGCCCCAGGGGCTGTGGACAAACCGGCTGGGGGACTGCC 742
QY 3772 CAGCCTTTTAC-----TGTTTAACTTATTTCACTGCTTTATATAAGGAAACACTAACA 3827
Db 743 CAGCCTCCACCTGTTTAACTTATTTCACTGCTTTATATAAGGAAACACTAACA 802
QY 3828 GA 3829
Db 803 AA 804

RESULT 8
US-09-966-262-48
; Sequence 48, Application US/09966262
; Publication No. US20030050461A1
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 87 Human Secreted Proteins
; FILE REFERENCE: PZ004P1
; CURRENT APPLICATION NUMBER: US/09/966,262
; PRIOR FILING DATE: 2001-10-01
; PRIOR APPLICATION NUMBER: US 09/154,707
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: PCT/US98/05311
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: US 60/041,277
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/042,344
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/041,276
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/041,281
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/048,094
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,350
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,188
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,135
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/050,937
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,187
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,099
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,352
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,186
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,069
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,095
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,131
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,096
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,355
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,160
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,351
; PRIOR FILING DATE: 1997-05-30

```


Query Match 13.7%; Score 528.4; DB 10; Length 851;
Best Local Similarity 83.3%; Pred. No. 3.6e-146; Indels 45; Gaps 7;
Matches 701; Conservative 0; Mismatches 96;
QY 2995 CACATGAAGACTCACAGTGTGAGAAACCTTCGGTGTGCGCGTGTCCATACGCGCTCT 3054
Db 1 CACATGAAGACACACAGTGTGAGAGCCCTTCGGTGTGCGCGTGTCTTATGCGCTCT 60
QY 3055 GCTCCTCTGATACCTGAAACGGCCACGCGGTCCACAGAGAGAAAGCCCTACAAG 3114
Db 61 CCTCATCTGATACCTGAAACGGCCACGCGGTCCATACAGAGAGAGCCCTACAAG 120
QY 3115 TGCCCCCTCTGTCCGTATGCTGTGGCAACCTTGCCCAACCTCAAGCGTATGTCGCATC 3174
Db 121 TGCCCCCTCTGTCCGTATGCTGTGGCAACCTTGCCCAACCTCAAGCGTATGTCGCATC 180
QY 3175 CACTCTGGTACAAACCTTTTCGGTGTAGCTTTCGACTGCACTACAGCTGCAACGAGTATG 3234
Db 181 CACTCTGGTACAAACCTTTTCGGTGTAGCTTTCGACTGCACTACAGCTGCAACGAGTATG 240
QY 3235 AACCTCAACGTCATATGCTGCCACACACGCGGAGAGAGCCCTTCGGCTGTGCCACCTGC 3294
Db 241 AACCTCAACGTCATATGCTGCCACACACGCGGAGAGAGCCCTTCGGCTGTGCCACCTGC 299
QY 3295 GCCTATACCAAGCCACTGGGCAACTCAAGCGTCAAGAGTGTGATGCCATGT 3354
Db 300 GCCTATACCAAGCCACTGGGCAACTCAAGCGTCAAGAGTGTGATGCCATGT 359
QY 3355 GGAGCAGAGGCGCTGTCTCTGCGCCCTGAGGCTGGCGCCACCTCATAGCCCAACC 3414
Db 360 GGAGCAGAGGCGCTGTCTCTGCGCCCTGAGGCTGGCGCCACCTCATAGCCCAACC 419
QY 3415 TCTGTTTGTAGCACTCGGGGTCCAGCAGCCCTGGGTGCTACTGGTAGCAGGGCTCTTCAT 3474
Db 420 TCTGTTTGTAGCTCTCGGGGCCACACAGCCCTGGGACTGTGGCAGCCGGGTGTCCAC 479
QY 3475 TCAGACTCACTTGAACCTAGTGTCTTTTACCTGGGGCTCTAGGAATTAGCCCTATG 3534
Db 480 ACAGACTCACTTGAACCTAGTGTCTTTTACCTGGGGCTCTAGGAATTAGCCCTATG 509
QY 3535 CCTCCTGCATTTTATACAAATGAACTAGAAACCAACCTTCCCTTCTCCCGCTGTCTCA 3594
Db 510 ---CCCATGTTTATACAGAGGACCAAGCCACCTT---TTTCTCCCGCTGTGCGCA 562
QY 3595 GGGGCTCCACACAGACTAACTAGGCACTATAGGACAGGCTGAATCCCATGGTCAGGG 3654
Db 563 GGGGCTCCACACAGACTAACTAGGCACTATAGGACAGGCTGAATCCCATGGTGGGG 622
QY 3655 GG-CCATATAGACACAGGGA-CTTGTCTTAGCTCACGTACAGATGAGCTAAGTGAATTAG 3712
Db 623 GGCCCATATGACACAGGACCTTTCCTTGTAGTGGGCACTTTCAGAGCTCAGTGAAGA 682
QY 3713 GGCCTTGAATTCACCGGCACTGTCTCCAGAGGCTATGATGAATGTTGGAG-CTGCC 3771
Db 683 GGCCTTGAATTCACCTCCACTGCCCCCAGGGGCTGTGGACAAACCGGCTGGGGACTGCC 742
QY 3772 CAGCCTTTTAC----TGTTTTAACTTATTTTCACTGCTTTTATAAAGGAAACACTAACA 3827

Db 743 CAGCTCCACCTGTATTAACTTTTCACTGCTTTTATAAAGGAAACACTAACA 802
QY 3828 GA 3829
Db 803 AA 804
RESULT 9
US-09-983-966-48
; Sequence 48, Application US/09983966
; Publication No. US20030060619A1
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 87 Human Secreted Proteins
; FILE REFERENCE: PZ04EP1
; CURRENT APPLICATION NUMBER: US/09/983,966
; PRIOR FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 09/154,707
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: PCT/US98/05311
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: US 60/041,277
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/042,344
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/041,276
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/041,281
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/048,094
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,350
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,188
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,135
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/050,937
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,187
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,099
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,352
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,186
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,069
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,095
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,131
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,096
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,355
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,160
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,351
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,154
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/054,804
; PRIOR FILING DATE: 1997-08-05
; PRIOR APPLICATION NUMBER: US 60/056,370
; PRIOR FILING DATE: 1997-08-19
; PRIOR APPLICATION NUMBER: US 60/060,862
; PRIOR FILING DATE: 1997-10-02
; NUMBER OF SEQ ID NOS: 343
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 48

```

; LENGTH: 851
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-983-966-48

Query Match      13.7%; Score 528.4; DB 10; Length 851;
Best Local Similarity 83.3%; Pred. No. 3.6e-146;
Matches 701; Conservative 0; Mismatches 96; Indels 45; Gaps 7;

QY 2995 CACATGAAGACTCACAGTGGTGAAGAAACCCCTTCGGCTGTGCCCTGTGCCATAGCCCTCT 3054
Db 1 CACATGAAGACACACAGTGGTGAAGAAACCCCTTCGGCTGTGCCCTGTGCCCTGTGCCCTCT 60

QY 3055 GCTCATCTGGATAAAGCTGAAGAGGACAGGCGCTGCACAGAGAGAAAGCCCTCAAG 3114
Db 61 CTTCTATCTGGATAAAGCTGAAGAGGACAGGCGCTGCACAGAGAGAAAGCCCTCAAG 120

QY 3115 TGCCCCCTCTGTCCGTATGCTGTGGCAACTTGGCCAACTCAAGCGTCAATGTCGCATC 3174
Db 121 TGCCCCCTCTGTCCGTATGCTGTGGCAACTTGGCCAACTCAAGCGTCAATGTCGCATC 180

QY 3175 CACTCTGGTGCAAAACCTTTTCGGTGTAGCTTTTGGCAACTAGTGCACAGATG 3234
Db 181 CACTCTGGTGCAAAACCTTTTCGGTGTAGCTTTTGGCAACTAGTGCACAGATG 240

QY 3235 RACCTCAAGCTCATATGTGCGACACAGCGCGAGAGCCCTTCGGCTGTGCCCTGTGCC 3294
Db 241 RACCTCAAGCTCATATGTGCGCGACACAGCGCGAGAGCCCTTCGGCTGTGCCCTGTGCC 299

QY 3295 GCTTATACACAGGCGCACTGGGACAACTACAAGCGTCTATGAAGGTGCATGGCCATGGT 3354
Db 300 GCTTATACACAGGCGCACTGGGACAACTACAAGCGCGCCACCAAGAGGTGCATGGCCACGGT 359

QY 3355 GAGCAGAGGCGCTGTCTCTCTCCCTGTAGGCGTGGCGCCACCTCATAGCCACCC 3414
Db 360 GAGCAGAGGCGCTGTCTCTCTCCCTGTAGGCGTGGCGCCACCTCATAGCCACCC 419

QY 3415 TCTGTTTTGAGCACTCGGGGTCAGCAGCGCTGGTGTCTTCTGTGTAGCAGGCTCTTCAT 3474
Db 420 TCTGTTTTGAGCTCTCGGGGTCAGCAGCGCTGGTGTCTTCTGTGTAGCAGGCTCTTCAC 479

QY 3475 TGAGCTACCTTGAACTAACTAGTGTCTTTTACCTGGGGCTCTAGGAATAGCCCTATG 3534
Db 480 ACAGACTCATCTGAACCTAGTGTCTTTTACCTGGGGCTCTAGGAATAGCCCTATG 509

QY 3535 CCTCTGCACTTTTATACAAATGAATAGAAACCACTTTTCTCCCTGCTGCTGCA 3594
Db 510 ---CCCATGTTTATACAGCGACAGAGCCACCTT---TTTCTCCCTGCTGCGCA 562

QY 3595 GGGGCTCCACACAGACTAAGCTAGGCACTATATGACAGCGCTGATCCCATGTCAGGG 3654
Db 563 GGGGCTCCACACAGACTAAGCTAGGCACTATATGACAGCGCGCCCAACCCCATGGGGGGG 622

QY 3655 GG-CCATATAGCCAGGGA-CTTGTCTTAGCTCAGTACAGATGAGCTAAGTGTATTAG 3712
Db 623 GGCCCATATGAGACAGGAGACCTTGTCTTGTAGCTGAGGCACTTCCAGGCTCAGTGAAG 682

QY 3713 GGCCTTGGATTACCCGCACTGTCTCCAGAGGCTATGATGAACTATGTTGGAG-CTGCC 3771
Db 683 GGCCTTGTATTACCTTCCACTGTGCGGCGCTGTGGAACAACCGCTGGGGGACTGCC 742

QY 3772 CAGCCTTTTAC-----TGTTTAACTATTTCAGTGTCTTATTAAGGAAACACTAACA 3827
Db 743 CAGCCTTCCACCTGTTTATTAACTATTTCAGTGTCTTATTAAGGAAACACTAACA 802

3828 GA 3829
803 AA 804

; Publication No. US20030018180A1
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 87 Human Secreted Proteins
; FILE REFERENCE: P2004P1
; CURRENT APPLICATION NUMBER: US/10/059,395
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: US/09/966,262
; PRIOR FILING DATE: 2001-10-01
; PRIOR APPLICATION NUMBER: US 09/154,707
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: PCT/US98/05311
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: US 60/041,277
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/042,344
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/041,276
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/041,281
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/048,094
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,350
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,188
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,135
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/050,937
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,187
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,099
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,352
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,186
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,069
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,095
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,131
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,096
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,355
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,160
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,351
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,154
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/054,804
; PRIOR FILING DATE: 1997-08-05
; PRIOR APPLICATION NUMBER: US 60/056,370
; PRIOR FILING DATE: 1997-08-19
; PRIOR APPLICATION NUMBER: US 60/060,862
; PRIOR FILING DATE: 1997-10-02
; NUMBER OF SEQ ID NOS: 343
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 48
; LENGTH: 851
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-059-395-48

```

Query Match 13.7%; Score 528.4; DB 13; Length 851;
Best Local Similarity 83.3%; Pred. No. 3.6e-146;
Matches 701; Conservative 0; Mismatches 96; Indels 45; Gaps 7;

2995 CACATGAAGACTCACAGTGTGAGAAACCTTCCGCTGTGCGCTGTCCATACGCTCT 3054
1 CACATGAAGACACACAGTGTGAGAAACCTTCCGCTGTGCGCTGTCCATACGCTCT 60
3055 GCTCATCTGATTAACCTGAAACGGCCACAGCGGTCCACACAGAGAGAAAGCCCTACAAG 3114
61 CCTCATCTGATTAACCTGAAACGGCCACAGCGGTCCATACAGAGAGAGAAAGCCCTACAAG 120
3115 TGCCCCCTCTGTCGATGCTGTGCGCAACCTGGCCAACTCAAGGCTATGTCGATC 3174
121 TGCCCCCTCTGTCGATGCTGTGCGCAACCTGGCCAACTCAAGGCTATGTCGATC 180
3175 CACTCTGGTGAACAACTTTTGGTGTAGCTTTGCAACTACAGCTGCAACAGAGATG 3234
181 CACTCTGGTGAACAACTTTTGGTGTAGCTTTGCAACTACAGCTGCAACAGAGATG 240
3235 AACCTCAAAAGCTCATATGCTGCGCACACAGGGCGAGAAAGCCCTGTGCGCACTGC 3294
241 AACCTCAAAAGCTCATATGCTGCGCACACAGGGCGAGAAAGCC- TTCCGCTGTGCGCACTGC 299
3295 GCCTATACACAGCGCACTGGGCAACTACAGGGTCTCAGAGGTGCAATGCCATGCT 3354
300 GCCTATACACAGCGCACTGGGCAACTACAGGGTCTCAGAGGTGCAATGCCATGCT 359
3355 GGAGCAGAGGGCTGTGCTCTTGCCCCCTGAGGCTGGGCCCCACCTCATAGCCCAACC 3414
360 GGAGCAGAGGGCTGTGCTCTTGCCCCCTGAGGCTGGGCCCCACCTCATAGCCCAACC 419
3415 TCTGTTTGTAGCACTCGGGGTCCAGAGCCCTGGGTGCTACTGTAGCAGGGCTCTTCAT 3474
420 TCTGTTTGTAGCACTCGGGGTCCAGAGCCCTGGGTGCTACTGTAGCAGGGCTCTTCAC 479
3475 TCAGACTCACCTTGAACTAACTAGTCTTTTACCTGGGGCTCTAGGAATTAGCCCTATG 3534
480 ACAGACTCATCTGAACTAGTCTTTCTTCT----- 509
3535 CCTCTGCAATTTTATACAAATGAACCTAGAAACCACTTTCCCTTTTCCCTGCTGTGCA 3594
510 ---CCCATGTTTATACAGAGCGGACCAAGACCACTT---TTTCTCCCTGCTGCGCA 562
3595 GGGGTCCACACAGACTAACTAGGCACTATATGACCGAGCTGATCCCATGCTCAGGG 3654
563 GGGGTCCACACAGACTAACTAGGCACTATATGACCGAGCGCAACCCCATGCGCGGGG 622
3655 GG-CCATATAGACACAGGGA-CTTGCTTTAGCTCAGTACCGTACAGATGAGCTAAGTATTAG 3712
623 GGGCCATATGACACAGGAGCTTGGCTTGTAGTGGCACTTCCAGGACTCAGTGAGAAG 682
3713 GGCCTTGGATTACCGGCACTGCTCCAGAGGCTATGATGAACTGGTTGGAG-CTGCC 3771
683 GGCCTTGGATTACCGGCACTGCTCCAGAGGCTATGATGAACTGGTTGGAG-CTGCC 742
3772 CAGCCTTTTAC----TGTTTTAACTTTTTCAGTGTCTTTTATAAAGGAAACACTAACA 3827
743 CAGCCTCCCACTGTTTATTAACTTTTTCAGTGTCTTTTATAAAGGAAACACTAACA 802
3828 GA 3829
803 AA 804

RESULT 11

US-10-143-090-48
; Sequence 48, Application US/10143090
; Publication No. US2003069406A1
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 87 Human Secreted Proteins
; FILE REFERENCE: P2004P1
; CURRENT APPLICATION NUMBER: US/10/143,090
; PRIOR FILING DATE: 2002-05-13
; PRIOR APPLICATION NUMBER: 09/154,707
; PRIOR FILING DATE: 1998-09-17

; NUMBER OF SEQ ID NOS: 343
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 48
; LENGTH: 851
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-143-090-48

Query Match 13.7%; Score 528.4; DB 15; Length 851;
Best Local Similarity 83.3%; Pred. No. 3.6e-146;
Matches 701; Conservative 0; Mismatches 96; Indels 45; Gaps 7;

QY 2995 CACATGAAGACTCACAGTGTGAGAAACCTTCCGCTGTGCGCTGTCCATACGCTCT 3054
Db 1 CACATGAAGACACACAGTGTGAGAAAGCCCTTCCGCTGTGCGCTGTCCATACGCTCT 60
QY 3055 GCTCATCTGATTAACCTGAAACGGCCACAGCGGTCCACACAGAGAGAAAGCCCTACAAG 3114
Db 61 CCTCATCTGATTAACCTGAAACGGCCACAGCGGTCCATACAGAGAGAGAAAGCCCTACAAG 120
QY 3115 TGCCCCCTCTGTCGATGCTGTGCGCAACCTGGCCAACTCAAGGCTATGTCGATC 3174
Db 121 TGCCCCCTCTGTCGATGCTGTGCGCAACCTGGCCAACTCAAGGCTATGTCGATC 180
QY 3175 CACTCTGGTGAACAACTTTTGGTGTAGCTTTGCAACTACAGCTGCAACAGAGATG 3234
Db 181 CACTCTGGTGAACAACTTTTGGTGTAGCTTTGCAACTACAGCTGCAACAGAGATG 240
QY 3235 AACCTCAAAAGCTCATATGCTGCGCACACAGGGCGAGAAAGCCCTGTGCGCACTGC 3294
Db 241 AACCTCAAAAGCTCATATGCTGCGCACACAGGGCGAGAAAGCC- TTCCGCTGTGCGCACTGC 299
QY 3295 GCCTATACACAGCGCACTGGGCAACTACAGGGTCTCAGAGGTGCAATGCCATGCT 3354
Db 300 GCCTATACACAGCGCACTGGGCAACTACAGGGTCTCAGAGGTGCAATGCCATGCT 359
QY 3355 GGAGCAGAGGGCTGTGCTCTTGCCCCCTGAGGCTGGGCCCCACCTCATAGCCCAACC 3414
Db 360 GGAGCAGAGGGCTGTGCTCTTGCCCCCTGAGGCTGGGCCCCACCTCATAGCCCAACC 419
QY 3415 TCTGTTTGTAGCACTCGGGGTCCAGAGCCCTGGGTGCTACTGTAGTACAGGGCTCTTCAT 3474
Db 420 TCTGTTTGTAGCTCTCGGGGCCACACAGCCCTGGGACTGCTGGCAGCGGGCTGTCCAC 479
QY 3475 TCAGACTCACCTTGAACTAACTAGTCTTTTACCTGGGGCTCTAGGAATTAGCCCTATG 3534
Db 480 ACAGACTCATCTGAACTAGTCTTTCTTCT----- 509
QY 3535 CCTCTGCAATTTTATACAAATGAACCTAGAAACCACTTTCCCTTTTCCCTGCTGTGCA 3594
Db 510 ---CCCATGTTTATACAGAGCGGACCAAGACCACTT---TTTCTCCCTGCTGCGCA 562
QY 3595 GGGGTCCACACAGACTAACTAGGCACTATATGACCGAGCTGATCCCATGCTCAGGG 3654
Db 563 GGGGTCCACACAGACTAACTAGGCACTATATGACCGAGCGCAACCCCATGCGCGGGG 622
QY 3655 GG-CCATATAGACACAGGGA-CTTGCTTTAGCTCAGTACCGTACAGATGAGCTAAGTATTAG 3712
Db 623 GGGCCATATGACACAGGAGCTTGGCTTGTAGTGGCACTTCCAGGACTCAGTGAGAAG 682
QY 3713 GGCCTTGGATTACCGGCACTGCTCCAGAGGCTATGATGAACTGGTTGGAG-CTGCC 3771
Db 683 GGCCTTGGATTACCGGCACTGCTCCAGAGGCTATGATGAACTGGTTGGAG-CTGCC 742
QY 3772 CAGCCTTTTAC----TGTTTTAACTTTTTCAGTGTCTTTTATAAAGGAAACACTAACA 3827
Db 743 CAGCCTCCCACTGTTTATTAACTTTTTCAGTGTCTTTTATAAAGGAAACACTAACA 802
QY 3828 GA 3829
Db 803 AA 804

Query Match 10.4%; Score 399.2; DB 9;
Best Local Similarity 93.0%; Pred. No. 8.5e-108;
Matches 452; Conservative 0; Mismatches 28; Indels 6

Fri May 14 09:10:45 2004

```
Query Match          9.8%; Score 376; DB 14; Length 1934;
Best Local Similarity 82.1%; Pred. No. 1.6e-100;
Matches 506; Conservative 20; Mismatches 70; Indels 20; Gaps 8;

QY 373 AGCTGCTGGATCAGTCCCTGGCGGCAGACACACTCTCTGGGATGGTACAGGGTGTGACACA 432
Db 1295 AGCTGCTGGATCAGTCCCTGGCGGCAGACACACTCTCTGGGATGGTACAGGGTGTGACACA 1354

QY 433 TGACGTGCATCATATTGTTTCAAGCCCGGAAACACAGTAGAGCTTACGCGAGAGGTG 492
Db 1355 TGACCTGCATCATATTGTTTCAAGCCCGGAAACACAGTAGAGCTTACGCGAGAGGTG 1414

QY 493 GCAAGAGGAACTGGAGGAGGCACTCTCCACGAGGGGCTGAAGACACCGGCAACAGTG 552
Db 1415 GCAGCGAAACCTAGAGGAGGTGCTCTTTACTAGGGGCTGAAGAAATGCGCAACAGCG 1474

QY 553 AC---AAAAGAGGCCAAGAGGGANTAGTGGTCAACCGGA-CCCTGCCATGTGGACTG 608
Db 1475 ACAAGAAGAGGAGGCCAAGCGAGACTAGCAGTCAATCCAGACCCCTGCCACACTAGAYTG 1534

QY 609 TTTTCTGAGCCCT-TGGACCGGAGACTGAGTTTGTCTTGTCTTGTAGCCTTAGCAGTG 667
Db 1535 TTTTGTAGCCCTTCCCGACCTGAGATGAGTTTGTCTTGTCTTGTAGCCTTAGCAGTG 1594

QY 668 GGTATGAGGTGTGCAGGGG--GCTGGGTGGCTTTCCTCAGCCCATPACAAAGAGGGCCCC 725
Db 1595 GGTATGAGGTGTGCAGGGGAGCTGGGTGGCTTTCAMTCGCCCATTCCAAAGAGGGYTCT 1654

QY 726 CCACCCCCCGGCGGAGCCTGGGAGGCTGTGCTCTTAAGCCTCTTACTCTC 785
Db 1655 CCYTCCACACTGCAG-----CCGGAGGYTGTGTCTTCCAGCCGCTYTG-YTC 1707

QY 786 CTTGGGCTCATCGACTATCGTTTCTGTGCTGTGCTGTGTTGGTGGAGGAAGGACTG 845
Db 1708 CTCGGYTATC---ACCGTTCTGTGCTGTGCTGTGTTGGTGGAGGAAGGACTG 1763

QY 846 GTAGTTCGTATTTTACTCTGTGAACACTTTATTTAAGGACATCTTTTATTGGCGC 905
Db 1764 GCGGTTCTGTTTTTACTCTGTGA--ACTTTATTTAAGGACATTTTATTATTGGCGC 1821

QY 906 TCTGTGACCCCTAGCCGCTTGGACCCGCTCTCTGTGTGTACACTTTCAGCAACACTTTT 965
Db 1822 TCCATGGCCCTCGGCCCTTGCACCCGCTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1881

QY 966 CAGACTAAAGGCCAAA 981
Db 1882 CAGACTAAAGGCCAAA 1897
```

Search completed: May 14, 2004, 06:56:11
Job time : 1483 secs

DATE TO BEAT =

Fri May 14 09:10:46 2004

us-09-673-994a-7.rst

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 13, 2004, 15:34:15 ; Search time 9218 Seconds
(without alignments)
12462.552 Million cell updates/sec

Title: US-09-673-994A-7
Perfect score: 3847
Sequence: 1 ggaaggttagactttatttc.....gaaaaa 3847

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:**
1: em_estba:**
2: em_esthum:**
3: em_estin:**
4: em_estmd:**
5: em_estov:**
6: em_estpl:**
7: em_estro:**
8: em_htc:**
9: gb_est1:**
10: gb_est2:**
11: gb_htc:**
12: gb_est3:**
13: gb_est4:**
14: gb_est5:**
15: em_estfun:**
16: em_estom:**
17: em_gss_hum:**
18: em_gss_inv:**
19: em_gss_pln:**
20: em_gss_vrt:**
21: em_gss_fun:**
22: em_gss_man:**
23: em_gss_mus:**
24: em_gss_pro:**
25: em_gss_rod:**
26: em_gss_phg:**
27: em_gss_vrl:**
28: gb_gss1:**
29: gb_gss2:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	827.2	21.5	1247	11 AK052459	AK052459 Mus muscu
2	732.4	19.0	906	9 AL534516	AL534516 AL534516
3	716.2	18.6	735	13 BQ749274	BQ749274 UI-M-FD0-
4	687.2	17.9	990	12 BI668143	BI668143 603295981

5	677	17.6	1069	13	BQ072985
6	673.4	17.5	1001	13	BQ056860
7	666.8	17.3	843	10	BF533875
8	644.2	16.7	929	13	BF533875
9	628	16.3	645	12	BI408818
10	623.6	16.2	694	10	AW990386
11	616.2	16.0	793	12	BG574616
12	613.8	16.0	628	10	AW536524
13	613.8	16.0	666	10	AW536842
14	613.8	16.0	692	12	BM114215
15	613.8	16.0	727	12	BM119847
16	612.8	15.9	616	12	BM245850
17	612.4	15.8	718	13	BX672987
18	607.4	15.7	652	12	BI903576
19	603.8	15.7	805	12	BI407148
20	602.8	15.7	629	10	AW539472
21	602.8	15.7	663	12	BI903756
22	587.8	15.3	632	14	CA450828
23	581.2	15.1	668	12	BG707470
24	576.8	15.0	581	13	C78804
25	567.2	14.7	703	9	AI327168
26	562.6	14.6	693	14	CB327866
27	561.2	14.6	566	10	AW538698
28	560.4	14.5	912	12	BG392489
29	559.4	14.5	905	9	AL565547
30	557.8	14.5	563	9	AI461955
31	557.6	14.5	939	10	BE908881
32	556	14.5	710	9	AI385652
33	551.8	14.3	567	10	AW538638
34	547	14.2	568	12	BI439920
35	546.8	14.2	550	10	AW544493
36	546.8	14.2	615	10	BB484959
37	544	14.1	705	12	BI407530
38	540.8	14.1	558	13	BQ921766
39	536.8	14.0	653	12	BM936232
40	536	13.9	663	13	BQ195083
41	529	13.8	587	13	BQ747039
42	528.6	13.7	607	13	BX670592
43	526.8	13.7	733	13	BQ203743
44	526.8	13.7	736	13	BQ208560
45	525.2	13.7	692	12	BI285247

ALIGNMENTS

RESULT 1
LOCUS AK052459 1247 bp mRNA linear HTC 20-SEP-2003
DEFINITION Mus musculus 13 days embryo lung cDNA, RIKEN full-length enriched library, clone:D430028M17 product:similar to CDNA FLJ32203 FIS, CLONE PLACE6003038, WEAKLY SIMILAR TO ZINC FINGER PROTEIN 84 [Homo sapiens], full insert sequence.

ACCESSION AK052459.1 GI:26342688
VERSION HTC; CAP trapper.
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1
AUTHORS Carninci, P. and Hayashizaki, Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636

REFERENCE 2
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, N., Komno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374

```

11042159
3
PUBMED
REFERENCE
AUTHORS
Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
Konno,H., Akiyama,J., Nishi,K., Kitsuai,T., Tashiro,H., Itoh,M.,
Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Haraada,A.,
Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
Fujiwaki,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
Yoneda,Y., Ishikawa,T., Osawa,K., Tanaka,T., Matsura,S., Kawai,J.,
Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530313
11076861
4
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5
The PANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 1247)
Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,M.,
Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kigawa,I., Kasukawa,T.,
Kato,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,P., Takaku-Akahira,S.,
Takeda,Y., Tanaka,T., Tomaru,A., Toyota,I., Yasunishi,A.,
Muramatsu,M. and Hayashizaki,Y.
Direct Submision
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan [E-mail:genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216]
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL:http://genome.gsc.riken.go.jp/
URL:http://fantom.gsc.riken.go.jp/
FEATURES
Location/Qualifiers
1..1247
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM,DB:D430028M17"
/db_xref="MGI:2421883"
/db_xref="taxon:10090"
/clone="D430028M17"
/tissue_type="lung"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stages="13 days embryo"
101..907
/Note="unnamed protein product; putative
similar to CDNA FLJ32203 FIS, CLONE PLACE6003038, WEAKLY
SIMILAR TO ZINC FINGER PROTEIN 84 [Homo sapiens]
(SPTFQ96ML3, evidence: PASTY, 75.2%ID, 58.9%length,
match=666)"
/codon_start=1
CDS

```

```

/protein_id="BAC35001.1"
/db_xref="GI:26342689"
/translation="MGTAWEDEGSTSTSLPSLVDEDSFDEGPGALVLESDLLLGQDLRF
EEEEDEGGDGHNDQMLMGFERDSQSGARPGLPYGLSDDESGGRLSAESVEEP
ARGFEASGKXPPRCARCPYASALDNLKRQVHTGKPYKPLCPYACGNLANLKR
HGRTHSGDKPPRCLNYSQNSMLKRLHMLRTGKPFRCATCAITTTGHWNDYKRHQ
KVHGGAGGPGRSAPGWAHPSPSVLSTRGPAALCATGSRALHSDSP"
1227..1232
polyA_signal
/Note="putative"
1247
polyA_site
/Note="putative"
ORIGIN
Query Match 21.5%; Score 827.2; DB 11; Length 1247;
Best Local Similarity 88.8%; Pred No. 2.9e-128;
Matches 895; Conservative 0; Mismatches 113; Indels 0; Gaps 0;
QY 2821 GGACAGGAACCTGGAGGGGTGAGGGCAGCAGGCTGGGAGTGCCTATGTGGGGCTGC 2880
DB 239 GAAGAGGAAGAGATGAAGGTGACGGCCACACACAGCAGCTCATGGGCTTTGAGAGAG 298
QY 2881 ATGCGAGGAGAGCTGGAGGGGTTCCTCCTGGGGACCCAGGGGCCCTGGTCAAAAGGC 2940
DB 299 TCTGAGGAGAGCTCTCAGGGGCCAGACTGGACTTCCCTATGGCTGAGTGACGACGAG 358
QY 2941 TTGCGCTGTAGTTTATGCCCTTTGCCACTACTACCCCAACCACTGGCTCGGCACATG 3000
DB 359 TCTGGGGCGGGCGGCACCTAAGTGGGAGAGTGAAGTTGAGAAACACGACCGGGGTCCA 418
QY 3001 AAGACTCAGAGGTGAGAAACCTTCGCTGTGCGCTGTGCCATACGCTCTGCTCAT 3060
DB 419 GGGAGGCGCAGTGGTGAGAAACCTTCGCTGTGCGCTGTGCCATACGCTCTGCTCAT 478
QY 3061 CTGGATAACCTGAAACGGCAGCAGCGCTGCCACAGAGAGAAACCTTACAAGTGCCCC 3120
DB 479 CTGGATAACCTGAAACGGCAGCAGCGCTGCCACAGAGAGAAACCTTACAAGTGCCCC 538
QY 3121 CTCTGTCCGTATGCCTGTGGCAACCTGCCACCTCAGCGTCATGTCGCTCCACTCT 3180
DB 539 CTCTGTCCGTATGCCTGTGGCAACCTGCCACCTCAGCGTCATGTCGCTCCACTCT 598
QY 3181 GGTGACAAACCTTTTCGGTGTAGCCTTTTGCAACTACAGCTGCAACACAGATGAACTC 3240
DB 599 GGTGACAAACCTTTTCGGTGTAGCCTTTTGCAACTACAGCTGCAACACAGATGAACTC 658
QY 3241 AATCGTCATATGTCGGACACACGGGGCGAAGCCCTTCGGCTGTGCCACCTGCGCCTAT 3300
DB 659 AATCGTCATATGTCGGACACACGGGGCGAAGCCCTTCGGCTGTGCCACCTGCGCCTAT 718
QY 3301 ACCACAGGCCTTGGGACAACTTACAAGCGTCATCAGAAGTGCATGGCCATGGTGAGCA 3360
DB 719 ACCACAGGCCTTGGGACAACTTACAAGCGTCATCAGAAGTGCATGGCCATGGTGAGCA 778
QY 3361 GGAGGCGCTGTCTCTCTGCCCTGAGGGCTGGGGCCCACTCATAGCCACCTCTGTT 3420
DB 779 GGAGGCGCTGTCTCTCTGCCCTGAGGGCTGGGGCCCACTCATAGCCACCTCTGTT 838
QY 3421 TTGAGCAGCTCGGGGTCCAGCAGCCCTGGGTGCTACTGTGTAGCAGGGCTCTTCAATTCAGAC 3480
DB 839 TTGAGCAGCTCGGGGTCCAGCAGCCCTGGGTGCTACTGTGTAGCAGGGCTCTTCAATTCAGAC 898
QY 3481 TCACCTTGAACCTAATAGTGTCTTTTACCTGGGGCTTAGGAAATTAGCCCTATGCCCTCT 3540
DB 899 TCACCTTGAACCTAATAGTGTCTTTTACCTGGGGCTTAGGAAATTAGCCCTATGCCCTCT 958
QY 3541 GCATTTTATACAACTGAACCTAGAAACCTTTTCCCTTTTCCCGCGGTGGTCAGGGGCT 3600
DB 959 GCATTTTATACAACTGAACCTAGAAACCTTTTCCCTTTTCCCGCGGTGGTCAGGGGCT 1018
QY 3601 CCACACAGACTAACCTAGGACCTATATGACACAGCCTGAATCCCATGGTGAAGGGGCCAT 3660
DB 1019 CCACACAGACTAACCTAGGACCTATATGACACAGCCTGAATCCCATGGTGAAGGGGCCAT 1078

```



```

QY 3661 ATAGACAGGGGACTGTCTTAGCTCACTACAGATGAGTAACTGATTAGGCGCTTGG 3720
Db 1079 ATAGACAGGGGACTGTCTTAGCTCACTACAGATGAGTAACTGATTAGGCGCTTGG 1138
QY 3721 ATTACCGCCACCTGCTCCAGAGGCTATCGATGAACCTGTTGGAGCTGCCCCAGCCTTTT 3780
Db 1139 ATTACCGCCACCTGCTCCAGAGGCTATCGATGAACCTGTTGGAGCTGCCCCAGCCTTTT 1198
QY 3781 ACTGTTTAACTTATTTCAGTCTGCTTTATTAATGAAGGAAACACTAACAG 3828
Db 1199 ACTGTTTAACTTATTTCAGTCTGCTTTATTAATGAAGGAAACACTAACAG 1246

RESULT 2
AL534516 906 bp mRNA linear EST 12-MAY-2003
LOCUS
DEFINITION
AL534516 Homo sapiens FETAL BRAIN Homo sapiens cDNA Clone
CS0DF004YC12 5-PRIME, mRNA sequence.
ACCESSION
VERSION
AL534516
SOURCE
EST.
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 906)
Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
COMMENT
On Feb 13, 2001 this sequence version replaced gi:12798009.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 1341.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DF004BB06QPI&cluster=1341.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paradise Avenue Genoscope sequence ID : CS0DF004BB06QPI.
FEATURES
Location/Qualifiers
source
1..906
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DF004YC12"
/tissue_type="FETAL BRAIN"
/dev_stage="fetal"
/clone_lib="Homo sapiens FETAL BRAIN"
/notes="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo (dT) primer. Five prime end
enriched, double-strand cDNA was digested with Not I and
cloned into the Not I and EcoRV sites of the pCMVSPORT 6
vector. Library was not normalized."

ORIGIN
Query Match 19.0%; Score 732.4; DB 9; Length 906;
Best Local Similarity 92.6%; Pred. No. 2.3e-112;
Matches 766; Conservative 2; Mismatches 59; Indels 0; Gaps 0;

QY 1917 AGTGATACCTAAGATTCCTTCGACGAGTGCTCTGGGCGCCCTGGTGTGGAGCGGATTT 1976
Db 80 AGTGATACCTAAGATTCCTTCGACGAGGAGCCCGGCGCCCTGGTGTGGAGCGGATTT 139
QY 1977 GCTACTAGCCCAAGATCTGGAGTTTGAAGAGAGAGAGAGAGATGAGTGAAGTGAAGCCCA 2036
Db 140 GCTACTAGCCCAAGATCTGGAGTTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 199
QY 2037 CAACGACGAGCTCATGGGCTTTGAGAGAGCTCTGAAGGAGCTCTCAGGGGGCCAGACC 2096
Db 200 CAGTGACGAGCTCATGGGCTTTGAGAGAGCTCTGAAGGAGCTCTCAGGGGGCCAGGCC 259

```

```

QY 2097 TGGACTTCCCTATGGGCTGAGTGACGACGAGTCTTGGGGGGCGCGCGCCTAAAGTGGGA 2156
Db 260 TGGGCTTCCCTATGGGCTGAGCGACGATGATCTTGGGGGGCGCGCGGCACTAAGTGGGA 319
QY 2157 GAGTGAAGTTAGAGAAACAGCAGGAGGGTCCAGGGAGGCCAGGGTGGAGAGGCCAGGCC 2216
Db 320 GAGTGAAGTTAGAGAGACAGCCAGGGGTCCAGGGAGGCCAGGGGTGGAGAGGCCAGGCC 379
QY 2217 AGCCTGTGAGTGTGGGGGGCGCACAGGTGAGGGGGCCGTGTTGTGGGGCAGAGAGGCG 2276
Db 380 AGCCTGCCAGCTGTGTGGGGGGCGCACAGGTGAGGGGGCCGTGTTGTGGGGCAGAGAGG 439
QY 2277 GGGTGGGGGGCCCCCGCTGCCCCCACCGTTACTGTACTCATGCCGGCTGTGCGCTTTGCT 2336
Db 440 GGGTGGGGGGCCCCCGCTGCCCCCACCGTTACTGTACTCATGCCGGCTGTGCGCTTTGCT 499
QY 2337 GTCCCACTACTCGAGGCCACCTGAAGCGGCACATGACAGACACACAGCGGGGAGAGCGCTT 2396
Db 500 GTCCCACTACTCGAGGCCACCTGAAGCGGCACATGACAGACACACAGCGGGGAGAGCGCTT 559
QY 2397 CGCTGTGGCGCTGCCCCATACGNTACGCCAGCTTGTGTCAACCTGACCGCACATACCG 2456
Db 560 CGCTGTGGCGCTGCCCCATACGNTACGCCAGCTTGTGTCAACCTGACCGCACATACCG 619
QY 2457 CACCCATACTGGCAGAGAGCCCTACCGTTGTCCCACTGCCCCCTTTGCTCGCAGCAGCT 2516
Db 620 CACCCATACTGGCAGAGAGCCCTACCGTTGTCCCACTGCCCCCTTTGCTCGCAGCAGCT 679
QY 2517 GGGCAACCTGAGGGCGCATCAGCGCACCCACAGGGGCTCCCACTCTCTCTGCCCCAAC 2576
Db 680 GGGCAACCTGAGGGCGCATCAGCGCACCCACAGGGGCTCCCACTCTCTCTGCCCCAAC 739
QY 2577 CTGTGGCTTTCGATGCTGTCTCCAGCAACCGGCTCCAGTCCCGCAGAGCAGCA 2636
Db 740 CTGTGGCTTTCGATGCTGTCTCCAGCAACCGGCTCCAGTCCCGCAGAGCAGCA 799
QY 2637 GGGGCAATGCCCCGACGATCAGAAATGCGGTGTATCTCTGCGAGACTTGAGTTCATGT 2696
Db 800 GGGGCGGGTGCCTGGCGGCGACCTGAAGATGCTGTCTCTCTCCAGATTGAGGCTCCATGT 859
QY 2697 GCACACAGTGTGTCAGCTTCTCTCCAGACTGTCGCGAGCTGCGGG 2743
Db 860 GCCACAGGTGTGCCAGTTTCTCTCCAGACTGTCGCGAGTGTGGAGTGTGGG 906

RESULT 3
BQ749274 735 bp mRNA linear EST 17-JUL-2002
LOCUS
DEFINITION
BQ749274 UI-M-FD0-byf-1-07-0-UI.r1 NIH_BMAP_F00 Mus musculus cDNA clone
IMAGE:5717526 5', mRNA sequence.
ACCESSION
VERSION
BQ749274.1 GI:21896061
KEYWORDS
EST.
MUS MUSCULUS (house mouse)
ORGANISM
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 735)
NIH-MGC http://img.nci.nih.gov/.
AUTHORS
National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE
Unpublished (1999)
JOURNAL
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
Tissue Procurement: Dr. James Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)

```

Seq primer: pyX-5, Location/Qualifiers 1. .735 /organism="Mus musculus" /mol_type="mRNA" /strain="C57BL/6" /db_xref="taxon:10090" /clone="IMAGE:5717526" /tissue_type="whole brain" /dev_stage="embryo 12.5 dpc" /lab_host="DH10B (T1 phage resistant)" /clone_lib="NIH_BMAP_FDO" /note="Organ: brain; Vector: pyX-Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured mRNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with Not I, and then cloned directionally into pyX-Asc vector. The library tag sequence located between the Not I site and the polyA tail, is TGAGAGGCC. This library was created for the University of Iowa Mouse Brain Molecular Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institutes of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."		18.6%; Score 716.2; DB 13; Length 735; Query Match 99.2%; Pred. No. 1.2e-109; Best Local Similarity 0; Mismatches 5; Indels 1; Gaps 1; Matches 729; Conservative	
ORIGIN			
QY	2838	GGGTGAGGGCAGCAGCGCTGGAGCTGCCATGTGTGGCGCTGCATCGCAGGAGAGCGCTGG	2897
Db	1	GGGTGAGGGCAGCAGCGCTGGAGCTGCCATGTGTGGCGCTGCATCGCAGGAGAGCGCTGG	60
QY	2898	AGGGGTGGCACT- GGGGAGCCCGAGGCGCTGGTGACAAAGCTTTGCTGTAGTTAT	2956
Db	61	AGGGGTGGCACTNGGGGAGCCCGAGGCGCTGGTGACAAAGCTTTGCTGTAGTTAT	120
QY	2957	GCCCTTTGGCACTCACTACCCCAACCACTGCTGCGCACATGAAGACTCAAGTGGTG	3016
Db	121	GCCCTTTGGCACTCACTACCCCAACCACTGCTGCGCACATGAAGACTCAAGTGGTG	180
QY	3017	AGAAACCTTCGGCTGTGCGCGCTGTCCATACGCTCTGCTCATCTGGATACCTGAAC	3076
Db	181	AGAAACCTTCGGCTGTGCGCGCTGTCCATACGCTCTGCTCATCTGGATACCTGAAC	240
QY	3077	GGCACAGCGCGTCCACAGAGGAGAAAGCCCTACAGTGCCTCCCTCTGCTGATGCT	3136
Db	241	GGCACAGCGCGTCCACAGAGGAGAAAGCCCTACAGTGCCTCCCTCTGCTGATGCT	300
QY	3137	GTGGCAACCTGGCCAACTCAAGCGTCACTGGTGGCATCCACTGTGGTGACAAACCTTTTC	3196
Db	301	GTGGCAACCTGGCCAACTCAAGCGTCACTGGTGGCATCCACTGTGGTGACAAACCTTTTC	360
QY	3197	GGTGTAGCTTTGCACTACAGTGGCAACAGAGTATGAACCTCAACGTCATATGCTGC	3256
Db	361	GGTGTAGCTTTGCACTACAGTGGCAACAGAGTATGAACCTCAACGTCATATGCTGC	420
QY	3257	GACACAGGGGAGAGCCCTTCGCTGTGCGACCTGCGCTATACCACAGGCCACTGGG	3316
Db	421	GACACAGGGGAGAGCCCTTCGCTGTGCGACCTGCGCTATACCACAGGCCACTGGG	480
QY	3317	ACAACTCAAGCGTCATCAGAAAGTGCATGCGCCATGTGTGAGCAGGAGGCGCTGCTCT	3376
Db	481	ACAACTCAAGCGTCATCAGAAAGTGCATGCGCCATGTGTGAGCAGGAGGCGCTGCTCT	540
QY	3377	CTGCCCCCTGAGGGCTGGGCCCCCACTCATAGCCCCACCCCTCTGTTTTCAGCACTCGGGGTC	3436
Db	541	CTGCCCCCTGAGGGCTGGGCCCCCACTCATAGCCCCACCCCTCTGTTTTCAGCACTCGGGGTC	600

QY	3437	CAGCAGCCCTGGGTGCTACTGCTAGCAGGGCTCTTCACTCAGACTCACCTTGAACCTAAT	3496
Db	601	CAGCAGCCCTGGGTGCTACTGCTAGCAGGGCTCTTCACTCAGACTCACCTTGAACCTAAT	660
QY	3497	AGGTTCTTTTACTGCGGCTCTAGGAATAGCCCTAGCTCTGCTGATTTTATACAAATG	3556
Db	661	AGGTTCTTTTACTGCGGCTCTAGGAATAGCCCTAGCTCTGCTGATTTTATACAAATG	720
QY	3557	AACTAGAAACCACT	3571
Db	721	AACTAGAAACCACTT	735
RESULT 4			
BI668143 990 bp mRNA linear EST 12-SEP-2001			
LOCUS 603295981F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5314868 5',			
DEFINITION mRNA sequence.			
ACCESSION BI668143			
VERSION BI668143.1 GI:15582376			
KEYWORDS EST.			
SOURCE Homo sapiens (human)			
ORGANISM Homo sapiens			
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
NIH-MGC http://mgc.nci.nih.gov/.			
National Institutes of Health, Mammalian Gene Collection (MGC)			
Unpublished (1999)			
Contact: Robert Strausberg, Ph.D.			
Email: cspbs@mail.nih.gov			
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.			
CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki			
Toshiyuki and Piero Carninci (RIKEN)			
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)			
DNA Sequencing by: Incyte Genomics, Inc.			
Clone distribution: MGC clone distribution information can be			
found through the I.M.A.G.E. Consortium/LLNL at:			
http://image.llnl.gov			
Plate: LUAM1798 row: b column: 21			
High quality sequence stop: 771.			
FEATURES			
Location/Qualifiers			
1. .990			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/clone="IMAGE:5314868"			
/tissue_type="hypothalamus"			
/lab_host="DH10B"			
/clone_lib="NIH_MGC_96"			
/note="Organ: brain; Vector: pBluescriptR (modified			
pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI			
(gtcgag); Oligo-dT primed using primer			
5'-TTTTTTTTTTTTTTVN-3', size-selected for average			
insert size 2.3 kb and normalized to ROT 5. This is a			
primary library enriched for full-length clones and			
constructed using the Cap-trapper method (Carninci, in			
preparation). Library constructed by M. Brownstein			
(NIH/NHGRI, National Institutes of Health). Note: this is			
a NIH_MGC Library."			
ORIGIN			
Query Match 17.9%; Score 687.2; DB 12; Length 990;			
Best Local Similarity 85.6%; Pred. No. 7.7e-105;			
Matches 847; Conservative 0; Mismatches 129; Indels 13; Gaps 7;			
QY	2352	CCACCTGAAGCGGCATGCAGACACACAGCGGGAGAGCGCTTCCGCTGTGGCGCTG	2411
Db	1	CCACCTGAAGCGGCATGCAGACACACAGCGGGAGAGCGCTTCCGCTGTGGCGCTG	60
QY	2412	CCCATACGCTCAGCCCACTTGGTCAACTGACGGGACATACCCCACTTACTGGGGA	2471
Db	61	CCCTAGCGCTCAGCCCACTTGGTCAACTGACGGGACATACCCCACTTACTGGGGA	120


```

481 ACCTGCGCGGCTGTGGACAAGAGCTGGAGGGGTGAGGGTAGTGGGTGGGAGCTGCC 540
2866 ATGTGTGGCGCTGATCGGAGAGAGGCTGGAGGGGTTGGCACTGGGGGAGCCCGAGGGC 2925
541 ATGTGTGGCGCTGTCATCGGAGAGAGGCTGGAGGGGTGCGCACTGGGGGGGCCAGGGC 600
2926 CTGTGTGACAAAGGCTTTGGCTGTAGTTTATGCGCCCTTTGCCACTCACTACCCCAACCA- 2984
601 CCCAGTGACAAAGGCTTTGGCTGTAGCTCTGCGCCCTTTGCCACTCACTATCCCAACCAC 660
2985 CTGTGTGGCGCATGAGAGCTCACTGTGTGAGAAACCTTCCGCTGTCGCGCTGTGCC 3044
661 CTGTGTGGCGCGCATGAGAGCACTGAGTGTGAGAAACCTTCCGCTGTCGCGCTGTGCC 720
3045 ATACGCGCTCTGCTCATCTGATTAACCTGAAACCGCACCGCGCTCCACACAGGA- GAAA 3103
721 TTATGCTCTGCTCATCTGATTAACCTGAAACCGCACCGCGCTCCATACAGGAGGA 780
3104 AGCCCTACAAAGTGCCTCC 3121
781 AGCCCTACCAATGCCCTCC 798

RESULT 6
BQ056860 1001 bp mRNA linear EST 29-MAR-2002
LOCUS AGENCOURT_6940531 NIH_MGC_99 Homo sapiens cDNA clone IMAGE:5809303
5', mRNA sequence.
ACCESSION BQ056860
VERSION BQ056860.1 GI:19816200
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE NIH-MGC http://mgi.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Lou Staudt
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2054 row: 1 column: 08
High quality sequence stop: 576.
FEATURES
Location/Qualifiers
1..1001
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5809303"
/tissue type="lymphoma, cell line"
/lab host="DH10B (phage-resistant)"
/clone lib="NIH_MGC_99"
note="Organ: lymph; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCACGAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library."
ORIGIN
Query Match 17.5%; Score 673.4; DB 13; Length 1001;
Best Local Similarity 91.8%; Pred. No. 1.5e-102;

```

```

Matches 765; Conservative 0; Mismatches 63; Indels 5; Gaps 5;
QY 2340 CCATATCTCAGCGCACTGAAGCGGCAATGCGAGACACACAGCGGGGAGAGCGGTTCGG 2399
Db 5 CCATATCTCAGCGCACTGAAGCGGCAATGCGAGACACACAGCGGGGAGAGCGGTTCGG 64
QY 2400 CTGTGGCGGCTGCCATACGCTCAGCCAGTTCGTCACCTGACGGGACATACCCCGCAC 2459
Db 65 CTGTGGCGGCTGCCCTTACGCTCAGCCAGTTCGTCACCTGACGGGACATACCCCGCAC 124
QY 2460 CCATATCTCAGCGAGAGCCCTACCGTTGTCCCACTGCCCCCTTTGGCTGCGAGCGCTGGG 2519
Db 125 CCACACTGGCGAGAGCCCTACCGCTGTCCCACTGCCCCCTTTGGCTGCGAGCGCTGGG 184
QY 2520 CAACCTGAGGGGCGATCAGCGCACCGACACAGGGGCTCCCACTCTCTCCCTGCCCAACTG 2579
Db 185 CAACCTGAGGGGCGATCAGCGTACCCACGAGGGGCCCCCACTCTCTCCCTGCCCAACTG 244
QY 2580 TGGCTTTTCGATGCTGTGCTCCACGACCAACCCCGGCTCCCACTGTCACAGCAGAGGGG 2639
Db 245 TGGCTTTTCGCTGTGTACTCCACGACCGCGGCTCCCACTGTCACAGCAGAGGGG 304
QY 2640 GACAAATGCCCCGAGATCAGAAATGCGCTGATCTCTGCGAGACTTGAGTCTTCATGTGCC 2699
Db 305 GCGGTGCCCCCGGCGACCTGAAGATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 364
QY 2700 ACCAGTGTGCTGCCAGTTTCTCTGCGAGACTGTGGGCGAGCTGCGGGGTGAAGGGGAGAGCTT 2759
Db 365 ACCAGTGTGCTGCCAGTTTCTCTGCGAGACTGTGGGCGAGCTGCGGGGTGAAGGGGAGGGCT 424
QY 2760 GTGTGGAACTGGATCCGAAACCACTGCGAGAGCTACTGTTCCTTTGGAGCTGCCGGGGCTG 2819
Db 425 GTGTGGAACTGGATCCGAAACCACTGCGAGAGCTACTGTTCCTTTGGAGCTGCCGGGGCTG 484
QY 2820 TGGACAGAACTGGAGGAGGTGAGGGGAGAGGCTGCGAGCTGCGGGGTGAGGGGAGAGCTG 2879
Db 485 TGGACAGAACTGGAGGAGGTGAGGGGTGAGGGGTGAGGGGTGAGGGGTGAGGGGTGAGGGGTG 544
QY 2880 CATGCGAGGAGAGGCTGGAGGGGTG-CACTGGGGGACCCCGAGGGGCTCGGTGACAAAG 2938
Db 545 CATGCGAGGAGAGGCTGGAGGGGTG-CACTGGGGGACCCCGAGGGGCTCGGTGACAAAG 604
QY 2939 GCTTTGCTGTAGTTTATGCGCCCTTTGCGAGCTCACTACCCCAACCCCTGCGGCGCTG 2998
Db 605 GCTTTGCTGTAGCTCTGCGCCCTTTGCGAGCTCACTATCCCAACCCCTGCGGCGCTG 664
QY 2999 TGAAGACTCACA-GTGGTGAGAAACCTTCCGCTGTGCGGCTGTGCCATAGCGCTTCTGCT 3057
Db 665 TGAAGACTCACA-GTGGTGAGAAACCTTCCGCTGTGCGGCTGTGCCATAGCGCTTCTGCT 724
QY 3058 CATCTGGATTAACCTGAAACGGGACAGCGGCTCCACACA-GGAGAAAGCCCTTCAAGT- 3115
Db 725 CATCTGGATTAACCTGAAACGGGACAGCGGCTCCATACAGGAGAAACCCCTTCAAGT 784
QY 3116 GCGCCCTCTGCTCGCTATGCTG-TGGCAACCTGGCCCAACCTCAAGCGCTCATG 3167
Db 785 GCGCCCTCTGCTCGCTATGCTG-TGGCAACCTGGCCCAACCTCAAGCGCTCATG 837

RESULT 7
BQ533875 843 bp mRNA linear EST 11-DEC-2000
LOCUS 602075222F1 NCI_CGAP_Li9 Mus musculus cDNA clone IMAGE:421245 5',
DEFINITION mRNA sequence.
ACCESSION BQ533875
VERSION BQ533875.1 GI:11621238
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 843)
NIH-MGC http://mgi.nci.nih.gov/.

```

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaps-r@mail.nih.gov
 Tissue Procurement: Jeffrey E. Green, M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Plate: L1AM9782 row: 1 column: 06
 High quality sequence stop: 817.

FEATURES
 source
 1..843
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clones="IMAGE:4212245"
 /lab_host="DH10B (T1 phage-resistant)"
 /clone_lib="NCI CGAP L49"
 /note="Organ: liver; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.9 kb. Constructed by Life Technologies. Note: this is a NCI CGAP Library."

ORIGIN
 Query Match 17.3%; Score 666.8; DB 10; Length 843;
 Best Local Similarity 95.8%; Pred. No. 2.1e-101;
 Matches 760; Conservative 0; Mismatches 23; Indels 10; Gaps 7;

QY 2330 CTTTCGTCTCCACTACTCGAGCCACCTGAAGCGCGCATGCGACACACAGCGGGGAGA 2389
 Db 1 CTTTCGTCTCCACTACTCGAGCCACCTGAAGCGCGCATGCGACACACAGCGGGGAGA 60

QY 2390 AGCGTTCCGTGTGGCGCTGCCATACGCTGAGCCAGTTCGTCACCTGACGGGAC 2449
 Db 61 AGCGTTCCGTGTGGCGCTGCCATACGCTGAGCCAGTTCGTCACCTGACGGGAC 120

QY 2450 ATACCGGACCCACTACTGGCGAGAGCCCTACCGCTGTGCTCCCTGCTCCCTTTCCCTGCA 2509
 Db 121 ATACCGGACCCACTACTGGCGAGAGCCCTACCGCTGTGCTCCCTGCTCCCTTTCCCTGCA 180

QY 2510 CGAGCTGGGCAACTGAGGGGGATGAGCGGACCCACAGGCGCTCCCACTCTCCCT 2569
 Db 181 CGAGCTGGGCAACTGAGGGGGATGAGCGGACCCACAGGCGCTCCCACTCTCCCT 240

QY 2570 GCCCAACTGTGGCTTTCGATGTGTGCTCCACAGCAACCCGCGCTCCCACTCCACAG 2629
 Db 241 GCCCAACTGTGGCTTTCGATGTGTGCTCCACAGCAACCCGCGCTCCCACTCCACAG 300

QY 2630 AGCAGAGGGGACAATGCCCGACATCAGAAATGCGCTGATCTGCGACACTTGATGTC 2699
 Db 301 AGCAGAGGGGACAATGCCCGACATCAGAAATGCGCTGATCTGCGACACTTGATGTC 360

QY 2690 TTCATGTCGCCACAGTGTGTCAGTTTCCTGCCAGACTGTGGGAGCTGGGGGTGAAG 2749
 Db 361 TTCATGTCGCCACAGTGTGTCAGTTTCCTGCCAGACTGTGGGAGCTGGGGGTGAAG 420

QY 2750 GGGAGAGCTTGTGTGGAATGTGATCCGAACCACTGCCAGACTACTGTTCCTTTGGACCT 2809
 Db 421 GGGAGAGCTTGTGTGGAATGTGATCCGAACCACTGCCAGACTACTGTTCCTTTGGACCT 480

QY 2810 GCCGGGGCTGTGACAGGAACTGGAGGAGGGGTGAGGGCAGCAGGCTGGAGCTGCCATGT 2869
 Db 481 GCCGGGGCTGTGACAGGAACTGGAGGAGGGGTGAGGGCAGCAGGCTGGAGCTGCCATGT 540

QY 2870 GTGGGCGCTGTCATGCGAGAGAGGCTGGAGGGGTTCCTCACTGGGGGA-CCCCAGGGCCCT 2928
 Db 541 GTGGGCGCTGTCATGCGAGAGAGGCTGGAGGGGTTCCTCACTGGGGGA-CCCCAGGGCCCT 600

QY 2929 GGTGCAAAAGGCTTGGCTGTGTAGTTATGTCCTTT-TCGCACTACTACCCCAACCACT 2987

Db 601 GTGTGCAAAAGGCTTGTGCTGTAGTTATGTCCTTTGTGCCACTACTACCCCAAC-ACC 658
 QY 2988 GCGTGGGCATGAAGACTCAGTGGTGAGAAACCTTCGCTGTGCCGTGTGCCATA 3047
 Db 659 TGGTGGGCATGAAGACTCAGTGGTGAGAAACCTTCCTCC---GTGTGCCGTGTGCATA 715
 QY 3048 CCGCTCTGCTCATCTGGATAACCTGAAACGGCAGCGCGTCCACAGAGAGAAAGCC 3107
 Db 716 CCGCTCTG-TCATCTGGATAACCTGAA-GGGACCATCGGT-CACACAGAGAGAAAGCC 772
 QY 3108 CTACAGTGGCCCC 3120
 Db 773 TACCAAGTGGCCC 785

RESULT 8
 BX453491/c
 LOCUS
 DEFINITION BX453491 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
 CSODF004YCL12 3-PRIME, mRNA sequence.
 BX453491
 ACCESSION BX453491.1 GI:31020508
 VERSION
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 CONTACT: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 1341.r For
 more information about this cluster, see
http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0BAI009ZF12_CS00862_2&cluster=1341.r.
 Contact : Feng Liang Email : fliang@lifetech.com URL :
<http://fulllength.invitrogen.com/InvitrogenCorporation1600>
 Paraday Avenue Genoscope sequence ID : CS0BAI009ZF12_CS00862_2.

FEATURES
 source
 1..929
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CSODF004YCL12"
 /tissue_type="FETAL BRAIN"
 /dev_stage="fetal"
 /clone_lib="Homo sapiens FETAL BRAIN"
 /notes="Organ: brain; Vector: pCMVSPORT_6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

ORIGIN
 Query Match 16.7%; Score 644.2; DB 13; Length 929;
 Best Local Similarity 83.5%; Pred. No. 1.2e-97;
 Matches 777; Conservative 0; Mismatches 149; Indels 4; Gaps 4;

QY 2072 AAGGAGACTCTCAGGGGGCCAGACCTCGGACTTCCCTATGGGCTGAGTACGACGAGTCTG 2131
 Db 929 AAGAAAGTCTTTGGGGGGGGGGGGTTCCTCCATAGTGTACACACTCGA-TTTG 871

QY 2132 GGGGGCGCGCGCACTAAGTGGCGGAGTCAAGTGGAGAACCCAGCCAGGGGTCCAGGGG 2191
 Db 870 GGGGGCGCGCGGGCCCTTAATCCGAGAGTAAATTTAGAAACTCACCCCGGGGTCCAGGG 811

QY 2192 AGGCCAGGGG-TGAGAGGCCAGGCCCGCCAGCTGTGCTG-TGGGGCCGACAGGTGA 2249

511 AACTACAGCGTCATCAGAAGTGCGATGGCCATGGTGGAGCAGGAGGGGCTTGGTCTCTCT 455

```

FEATURES
source
Location/Qualifiers
1. .793
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4705610"
/tissue_type="mammary adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 87"
/note="Organ: breast; Vector: pCMV-Sport6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.383 kb. Library enriched for

```


full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

ORIGIN
Query Match 16.0%; Score 616.2; DB 12; Length 793;
Best Local Similarity 89.9%; Pred. No. 6e-93;
Matches 716; Conservative 0; Mismatches 74; Indels 6; Gaps 5;
QY 2137 GSCCGCGCTAAGTTCGCGAGAGTGAAGTTGAGGAACAGCAGCGGTTCACAGGGAGGCC 2196
DB 1 GACCGCGTGGTTCGACCCACGCGTTCGAGTTCGAGGAGCCAGCAGGG--TCCAGGGAGGCC 58
QY 2197 AGGGGTGAGAGCCAGCCAGCCCTGTTCAGCTGTGTGTGGGGGCCGACAGGTGAGGGCCG 2256
DB 59 AGGTGTGAGAGCCAGCCAGCCCTGCCAGTGTGTGGGGGCCGACAGGTGAGGGCCG 118
QY 2257 TGTGTGGGGAGGAGGGGGGGTGGGGGGCCCGCTGCCCGTCCCGCCACGGTACTGTACTCA 2316
DB 119 TGTGTGGGGAGGAGGGGGGGTGGGGGGCCCGTGTGCCCGCCAGCTACTGTACTCA 178
QY 2317 TCCCGGCTGTGC-GCTTTCGTTCCTCACTACTCGAGCCACCTGTGAAGCGGCACATGCAGAC 2375
DB 179 TCCCGGCTGTGCCTTTCGTTCCTCACTACTCGAGCCACCTGTGAAGCGGCACATGCAGAC 238
QY 2376 ACACAGCGGGAGAGCCGCTTCGCTGTGGCGCTGCCCATACGCTCAGCCAGTTCGT 2435
DB 239 ACACAGCGGAGAGAGCCGCTTCGCTGTGGCGCTGCCCATACGCTCAGCCAGTTCGT 298
QY 2436 CAACCTGACCGACATACCCGACCCACATCTGCGAGAGCCCTACCGTTCGCCACTG 2495
DB 299 CAACCTGACCGACATACCCGACCCACATCTGCGAGAGCCCTACCGTTCGCCACTG 358
QY 2496 CCGCTTTCCTGACAGAGAGAGGGGCAATGCCCCGACATCAGCAATGCGCTGATCT 2675
DB 359 CCGCTTTCCTGACAGAGAGAGGGGCGGTGCCCCGACATCAGCAATGCGCTGATCT 538
QY 2676 GCGAGA-CTTGAGCTTCATGTCACACAGTGTGTCAGTTCCTGCCAGACTGTGGGC 2734
DB 539 TCCAGATTTGAGCTTCATGTCACACAGTGTGTCAGTTCCTGCCAGACTGTGGGC 598
QY 2735 AGCTGCGGGGTGAAGGGAGAGCTTGTGTGAATCTGGATCCGAACTCCAGAGCTAC 2794
DB 599 AGCTGCGGGGTGAAGGGAGAGGCTTGTGTGAATCTGGATCCGAACTCCAGAGCTAC 657
QY 2795 TGTTCCTTGGACCTGCGGGGCTG-TGGACAGAACTGGAGAGGCTGAGGCGAGCAGG 2853
DB 658 TATTCCTTGGACCTGCGGGGCTGTTGGACAGAGCTGGAGAGGCTGAGGCTAGTGG 717
QY 2854 CTGGAGCTGCCATGTGTGGCGCTGCATGCGAGGAGAGCTGGAGGGTTCACCTGGG 2913
DB 718 CTGGAGCTGCCATGTGTGGCGCTGCATGCGAGGAGAGCTGGAGGGTTCACCTGGG 777
QY 2914 GGACCCAGGCGCTG 2929
DB 778 GGGCCCCAGGCCCGAG 793

RESULT=12
AW536524/c 628 bp mRNA linear EST 31-AUG-2000
LOCUS AW536524
DEFINITION G0105E04-3 NIA Mouse E7.5 Embryonic Portion cDNA Library Mus
musculus cDNA clone G0105E04 3', mRNA sequence.
ACCESSION AW536524
VERSION AW536524.1 GI:7178941
KEYWORDS EST.

SOURCE
ORGANISM Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 628)
AUTHORS Tanaka,T.S., Jaradat,S.A., Lim,M.K., Kargul,G.J., Wang,X.,
Grabovac,M.J., Pantano,S., Sano,Y., Piao,Y., Nagatsuma,R., Doi,H.,
Wood,W.H., III, Becker,K.G. and Ko,M.S.H.
Genome-wide expression profiling of mid-gestation placenta and
embryo using a 15,000 mouse developmental cDNA microarray
Proc. Natl. Acad. Sci. U.S.A. 97 (16), 9127-9132 (2000)
MEDLINE 20381348
PUBMED 10922068
COMMENT Contact: George J. Kargul
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdna@leuun.grc.nia.nih.gov
Plate: G0105 row: E column: 04
Seq primer: -21M13 Forward
High quality sequence stop: 628
POLYA-Yes.
Location/Qualifiers
1..628
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="nia:EST:G0105E04-3"
/db_xref="taxon:10090"
/clone="G0105E04"
/sex="unknown"
/dev_stage="7.5dpc Embryo"
/lab_host="DH10B"
/clone_lib="NIA Mouse E7.5 Embryonic Portion cDNA Library"
/notes="Vector: pSPORT1 (Gibco/BRL Life Technology);
Site 1: SalI; Site 2: NotI. Total RNAs were extracted from
6 Embryo. The double-stranded cDNA was synthesized by
Gibco's kit with an oligo(dT) primer [NotI primer-adaptor
from Gibco/BRL]
[5'-pGACTAGTTCTAGATCGAGCGCGCCCTTTTCTTTT-3']
from 0.5ug of mRNA. The double-stranded cDNAs were
treated with T4 DNA polymerase and purified by
ethanol-precipitation. The cDNAs were ligated to
lone-linker JJ-Sal3 (include SalI sequence). The cDNAs
were purified by phenol/chloroform and separated from
free linkers by Centricon 100. Then, cDNAs were amplified
by long-range high fidelity PCR using Takara's Ex Taq
polymerase. Then, the cDNAs were purified by
phenol/chloroform and by Centricon 100. The cDNAs were
digested with SalI and NotI enzymes. Then, the cDNAs were
size selected by Gibco's Size Fractionation Column. The
cDNAs were cloned into SalI/NotI site of pSPORT1 plasmid
vector. The DH10B E. coli host was transformed with the
ligation mixture by chemical method. The library was
constructed by Xiaohong Wang and Minoru S. H. Ko."

ORIGIN
Query Match 16.0%; Score 613.8; DB 10; Length 628;
Best Local Similarity 99.7%; Pred. No. 1.7e-92;
Matches 615; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 373 AGCTGCTGGATCAGTTCCTGCGCCAGACACTTCTGGGATGTTACAGGTTGACACAA 432
DB 617 AGCTGCTGGATCAGTTCCTGCGCCAGACACTTCTGGGATGTTACAGGTTGACACAA 558
QY 433 TGACGTGATCATCATTTGTTTCAAGCCCCGAAACACAGTAGAGCTTCAGCAGAGAGTG 492
DB 557 TGACGTGATCATCATTTGTTTCAAGCCCCGAAACACAGTAGAGCTTCAGCAGAGAGTG 498
QY 493 GCAAGAGGAAACTGGAGGAGGAGCTGTCCAGGAGGGGGCTGGAAGACACCCGGCAACAGTG 552
DB 497 GCAAGAGGAAACTGGAGGAGGAGCTGTCCAGGAGGGGGCTGGAAGACACCCGGCAACAGTG 438


```

RESULT 14
BM114215/c
LOCUS
DEFINITION
L0801D02-3 NIA Mouse Newborn Brain cDNA Library Mus musculus cDNA
clone L0801D02 3', mRNA sequence.
ACCESSION
BM114215
VERSION
BM114215.1 GI:17077233
KEYWORDS
EST.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
REFERENCE
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 692)
Piao, Y., Kargul, G.J., Dudekula, D.B., Qian, Y., Pantano, S., Lim, M.K.
and Ko, M.S.H.
SYSTEMATIC ANALYSES OF NIA Mouse Newborn Brain cDNA Library
UNPUBLISHED (2001)
CONTACT: Dawood B. Dudekula
LABORATORY OF GENETICS
NATIONAL INSTITUTE ON AGING/NATIONAL INSTITUTES OF HEALTH
333 CASSELL DRIVE, SUITE 4000, BALTIMORE, MD 21224-6820, USA
Email: cdna@lgsun.grc.nia.nih.gov
Plate: L0801 row: D column: 02
Seq primer: -21M13 Forward
High quality sequence stop: 692
POLYA=Yes.
FEATURES
Location/Qualifiers
1..692
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="niaEST:L0801D02-3"
/db_xref="taxon:10090"
/clone="L0801D02"
/tissue_type="Newborn Brain"
/lab_host="DHI08"
/clone_stage="Newborn"
/notes="Organ: Brain; Vector: pSPORT1 (Invitrogen); Site: 1;
SalI; Site 2: NotI; Mouse cDNA project by the Laboratory
of Genetics, National Institute on Aging (NIA), Intramural
Research Program, NIH (http://lgsun.grc.nia.nih.gov/cDNA).
Double-stranded cDNAs were synthesized with an Oligo(dT)
primer [Invitrogen: 5'-
pGACTAGTTCTAGATCGGACGAGCGCCGCCCTTTT-3'] from 48
microgram of total RNA, treated with T4 DNA polymerase,
and purified by ethanol-precipitation. The cDNAs were
ligated to Loxe-linker LL-Sal3 (Ref. Development
127:1737-1749 (2000) [PMID:10725249]), purified by
phenol/chloroform, and separated from free linkers by
centrifugation 100. Then, the cDNAs were digested with SalI and
NotI enzymes, and cloned into SalI/NotI site of pSPORT1
plasmid vector. The DH10B E. Coli host was transformed
with ligation mixture by the standard chemical method. The
average insert size is about 1.9 kb. The library was
constructed by Yulan Piao (NIA)."
ORIGIN
Query Match 16.0%; Score 613.8; DB 12; Length 692;
Best Local Similarity 99.7%; Pred. No. 1.6e-92;
Matches 615; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 373 AGCTGCTGATCAGTCTGCTGGCCGACACCTCTCTGGGATGGTACAGGTTGTGACACA 432
DB 617 AGCTGCTGATCAGTCTGCTGGCCGACACCTCTCTGGGATGGTACAGGTTGTGACACA 558
QY 433 TGAGTGCATCATCTTTTTCAGGCCCGAACAACACAGTAGAGCTTCAGCAGAGAGTG 492
DB 557 TGAGTGCATCATCTTTTTCAGGCCCGAACAACACAGTAGAGCTTCAGCAGAGAGTG 498
QY 493 GCAAGAGAAACTGGAGGAGGCACTGTCCAGGGGGGCTGAGACACCGGCAACAGTG 552

```

```

497 GCAAGAGAAACTGGAGGAGGCACTGTCCACGAGGGGCTGAAGACACCCGGCAACAGTG 438
553 ACAGAGAGAGGCGCAAGAGGATTAAGTGTCAACCGGACCTCTGCCCATGTGGAATCTTTT 612
437 ACAGAGAGAGGCGCAAGAGGATTAAGTGTCAACCGGACCTCTGCCCATGTGGAATCTTTT 378
613 CTGAGCCCTTTGGACCCGAGACTGAGTTTTGTCTCTGCTCTCTTAAGCCTTCTAGCAGTGGGTAT 672
377 CTGAGCCCTTTGGACCCGAGACTGAGTTTTGTCTCTGCTCTCTTAAGCCTTCTAGCAGTGGGTAT 318
732 GAGGTGTGAGGGGGTGGGTGGCTTTCTCAGCCCATTAACAAGAGGGGCCCCCAGCCCC 732
317 GAGGTGTGAGGGGGTGGGTGGCTTTCTCAGCCCATTAACAAGAGGGGCCCCCAGCCCC 258
733 CCCACACCGGCGAGCCTTGGAGGCTCTGCTCTCTCTTAAGCCTTCTCTTACTCTCTTGGGC 792
257 CCCACACCGGCGAGCCTTGGAGGCTCTGCTCTCTCTTAAGCCTTCTCTTACTCTCTTGGGC 198
793 TCATCGACTATCGGTTCTGTGCTGCTGCTGCTCTCTTAAGCCTTCTCTTACTCTCTTGGGC 852
197 TCATCGACTATCGGTTCTGTGCTGCTGCTGCTCTCTTAAGCCTTCTCTTACTCTCTTGGGC 138
853 TGATTTTACTCTGTGAACACTTTATTAAGGACATTTCTTTTATTGGCGGCTCTGTGA 912
137 TGATTTTACTCTGTGAACACTTTATTAAGGACATTTCTTTTATTGGCGGCTCTGTGA 78
913 CCCTAGCCGCTTGCACCCGCTCTCTGTTGTACACTTTCAAGCAACACTTTTTCAGACTA 972
77 CCCTAGCCGCTTGCACCCGCTCTCTGTTGTACACTTTCAAGCAACACTTTTTCAGACTA 18
973 AAGGCCAAACAAAAGCT 989
17 AAGGCCAAACAAAAGCT 1
DB
RESULT 15
BM119847/c
LOCUS
DEFINITION
L0931C10-3 NIA Mouse Newborn Kidney cDNA Library (Long) Mus
musculus cDNA clone NIA:L0931C10 IMAGE:30003009 3', mRNA sequence.
ACCESSION
BM119847
VERSION
BM119847.2 GI:31556676
KEYWORDS
EST.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 727)
Piao, Y., Ko, N.T., Lim, M.K. and Ko, M.S.H.
CONSTRUCTION OF LONG-TRANSCRIPT ENRICHED cDNA LIBRARIES FROM
SUBMICROGRAM AMOUNTS OF TOTAL RNAs BY A UNIVERSAL PCR AMPLIFICATION
METHOD
Genome Res. 11 (9), 1553-1558 (2001)
21429098
11544199
On Nov 26, 2001 this sequence version replaced gi:17087873.
Contact: Dawood B. Dudekula
LABORATORY OF GENETICS
NATIONAL INSTITUTE ON AGING/NATIONAL INSTITUTES OF HEALTH
333 CASSELL DRIVE, SUITE 4000, BALTIMORE, MD 21224-6820, USA
Email: cdna@lgsun.grc.nia.nih.gov
Plate: L0931 row: C column: 10
Seq primer: -21M13 Forward
High quality sequence stop: 727
POLYA=Yes.
FEATURES
Location/Qualifiers
1..727
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="niaEST:L0931C10-3"
/db_xref="taxon:10090"
/clone="NIA:L0931C10 IMAGE:30003009"

```

/tissue_type="Newborn Kidney"
/dev_stage="Newborn"
/lab_host="DH10B"
/clone_lib="NIA Mouse Newborn Kidney cDNA Library (Long)"
/note="Organ: kidney; Vector: pSPORT1 (Invitrogen);
Site 1: SalI; Site 2: NotI; Mouse cDNA project by the
Laboratory of Genetics, National Institute on Aging (NIA),
Intramural Research Program, NIH
(http://lgsun.grc.nia.nih.gov/cDNA). This is a
long-transcript enriched cDNA library (Ref. Genome Res.
11:1553-1558 (2001). [PMID:11544199]). In
brief double-stranded cDNAs were synthesized with an
Oligo(dT) primer [Invitrogen: 5'-
pGATAGTCTAGATCGAGCGCGGCCCTTTT-3'] from 26
microgram of total RNA, treated with 14 DNA polymerase,
and purified by ethanol-precipitation. The cDNAs were
ligated to Lene-linker LL-Sal4, purified by
phenol/chloroform, and separated from free linkers by
Centricon 100. Then, the cDNAs were amplified by
long-range high fidelity PCR using Ex Tag polymerase
(Takara) with a primer Sal4-S. The products were purified
by phenol/chloroform and Centricon 100. The cDNAs were
digested with SalI and NotI enzymes, and cloned into
SalI/NotI site of pSPORT1 plasmid vector. The DH10B E.
coli host was transformed with ligation mixture by the
standard chemical method. The average insert size is about
3.0 kb. The library was constructed by Yulan Piao (NIA)."

ORIGIN

Query Match 16.0%; Score 613.8; DB 12; Length 727;
Best Local Similarity 99.7%; Pred. No. 1.6e-92;
Matches 615; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY	373	AGCTGTGATCAGTCCCTGGCCAGACACTCTCGGGATGGTACAGGGTGTGACACA	432
Db	617	AGCTGTGATCAGTCCCTGGCCAGACACTCTCGGGATGGTACAGGGTGTGACACA	558
QY	433	TGACGTGCATCATCTTTGTTTCAAGCCCCGAAACACAGTAGAGCTTCAGGCAGAGAGTG	492
Db	557	TGACGTGCATCATCTTTGTTTCAAGCCCCGAAACACAGTAGAGCTTCAGGCAGAGAGTG	498
QY	493	GCAAGAGGAACTGGAGGAGCACTGTCCAGGAGGGGGCTGAAGACACCGGCACAGTG	552
Db	497	GCAAGAGGAACTGGAGGAGCACTGTCCAGGAGGGGGCTGAAGACACCGGCACAGTG	438
QY	553	ACAAAAGAGGCAAGAGGAGTATGTGTCAACCGGACCCCTGCCCATGTGACTGTGTTT	612
Db	437	ACAAAAGAGGCAAGAGGAGTATGTGTCAACCGGACCCCTGCCCATGTGACTGTGTTT	378
QY	613	CTGAGCCCTTGGACCCGAGACTGAGTTTGTCTTGTCTTGTAGCCTTAGCAGTGGGTAT	672
Db	377	CTGAGCCCTTGGACCCGAGACTGAGTTTGTCTTGTCTTGTAGCCTTAGCAGTGGGTAT	318
QY	673	GAGGTGTGAGGGGGCTGGGTGCTTCTCAGCCCATTTACAAAGAGGGCCCCACACCC	732
Db	317	GAGGTGTGAGGGGGCTGGGTGCTTCTCAGCCCATTTACAAAGAGGGCCCCACACCC	258
QY	733	CCCCACGCGCAGCCTGGGAGGCTCTGCTGCTCTCTTAAGCCTCTTACTCTCTCTTGGGC	792
Db	257	CCCCACGCGCAGCCTGGGAGGCTCTGCTGCTCTCTTAAGCCTCTTACTCTCTCTTGGGC	198
QY	793	TCATCGACTATCGGTTCTGTGCTGCTCTGTTGTTGGAGGAGGACTGGTAGTTTC	852
Db	197	TCATCGACTATCGGTTCTGTGCTGCTCTGTTGTTGGAGGAGGACTGGTAGTTTC	138
QY	853	TGATTTTACTCTGTGAACACTTTATTTAAGGACATTTCTTTTATTTGGCGGCTCTGTGA	912
Db	137	TGATTTTACTCTGTGAACACTTTATTTAAGGACATTTCTTTTATTTGGCGGCTCTGTGA	78
QY	913	CCCTAGCCGCTTGACCCGCTCTCTGTTGTGTACACTTTCAAGCAACACTTTTTCAGACTA	972
Db	77	CCCTAGCCGCTTGACCCGCTCTCTGTTGTGTACACTTTCAAGCAACACTTTTTCAGACTA	18

QY 973 AAGCCCAACAAAAGCT 989
Db 17 AAGCCCAACAAAAGCT 1

Search completed: May 14, 2004, 05:45:48
Job time : 9239 secs